ence Library

AUG 2 4 1945

Vol. 35, No. 4 August, 1945

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, ON LEAVE
FRANCIS W. IRWIN, ACTING EDITOR
UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CHARLES W. BRAY, *Princeton University;* ELMER K. CULLER, *University of Rochester;* CLARENCE H. GRAHAM, *Brown University;* JOY P. GUILFORD, *University of Southern California;* FRANCIS W. IRWIN, *University of Pennsylvania;* DONALD G. MARQUIS, *Yale University;* ARTHUR W. MELTON, *University of Missouri;* CARROLL C. PRATT, *Rutgers University.*

CONTENTS

An Experimental Test of the Continuity and Non-Continuity Theories of Discrimination Learning: K. W. SPENCE 253

The Effect of Electroshock Convulsions on the Maze Habit in the White Rat: C. P. DUNCAN 267

Heart Rate and Frequency of Blinking as Indices of Visual Efficiency: M. E. BITTERMAN 279

Habit Strength as a Function of the Pattern of Reinforcement: O. H. MOWRER AND HELEN JONES 293

Effect of Various Concomitant Activities upon the Maintenance of a Set: DOROTHY RETHLINGSHAFER 312

Standardization of Two Tests of Equilibrium: the Railwalking Test and the Ataxiagraph: M. B. FISHER, J. E. BIRREN, AND A. L. LEGGETT 321

The Construction of a Punch Maze: W. A. BOUSFIELD AND V. NOWLIS 330

PUBLISHED BI-MONTHLY,
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

PUBLICATIONS OF

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.
Subscription: $5.50 (Foreign, $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

JOHN E. ANDERSON, *Editor*
University of Minnesota

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears monthly (10 issues), the annual volume comprising about 665 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor, on Leave,* FRANCIS W. IRWIN, *Acting Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears bi-monthly (beginning 1944), one volume per year, each volume of six numbers containing about 520 pages.
Subscription: $7.00 (Foreign, $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 350 pages.
Subscription: $6.00 per volume (Foreign, $6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 400 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.
Subscription: $5.00 (Foreign, $5.25). Single copies, $1.50.

JOURNAL OF APPLIED PSYCHOLOGY

DONALD G. PATERSON, *Editor*
University of Minnesota

Covers the applications of psychology in business, industry, education, etc. Appears bi-monthly, February, April, June, August, October, and December.
Subscription: $6.00. Single copies, $1.25.

Subscriptions, orders, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Journal of
Experimental Psychology

VOL. 35, NO. 4 AUGUST, 1945

AN EXPERIMENTAL TEST OF THE CONTINUITY AND NON-CONTINUITY THEORIES OF DISCRIMINATION LEARNING

BY KENNETH W. SPENCE

State University of Iowa

I. INTRODUCTION

Beginning with the studies of Hamilton (3) and Yerkes (19), a number of investigations involving different types of situations and a variety of animal Ss have shown that the pre-solution period of learning is an organized, lawful, systematic process. Thus, in a series of experiments with rats involving soluble and unsoluble discrimination problems, Krechevsky (6, 7) demonstrated that this period was marked by the occurrence of systematic response tendencies or, as he termed them, hypotheses. These hypotheses, precisely defined by Krechevsky in statistical terms as responses which occurred within a block of trials with a frequency beyond the limits of chance expectancy, consisted in his experiments primarily of various types of spatial responses, e.g., left or right position habits, an alternating habit, etc. Employing the same statistical criterion, Spence found similar types of systematic response tendencies occurring among chimpanzees during the pre-solution period of multiple choice problems (17).

Contrary to the beliefs of certain recent writers (5), then, there has been no disagreement concerning the behavioral facts. The pre-solution behavior of animals in such experimental situations as the multiple choice and discrimination problems is not a random, chance affair in the sense of being undetermined or unlawful. Instead it is to be described as consisting of systematic reactions or 'hypotheses,' if one defines this latter term in the statistical manner described above. Agreement ceases, however, over attempts at further interpretation of the phenomena.

From a suggestion originally made by Lashley (**10**) has developed one theoretical interpretation of these phenomena which has come to be known as the non-continuity theory. Lashley's suggestion was that the systematic response tendencies, such as response to position in the discrimination situation, represent attempted solutions of the problem on the part of the animal. Furthermore, Lashley intimated that during the part of the practice (learning) period in which these attempts at solution were being made no association would be formed with respect to the cue stimuli. Krechevsky subsequently formulated this interpretation in the following manner:

". . . Once an animal is immersed in a given problem-situation the animal selects out of the welter of possible stimuli certain sets of discriminanda to which he reacts. Each time (while 'paying attention to' this particular set of discriminanda) he makes what proves to be a 'correct' response, he learns (wrongly perhaps) something about the significance *of this particular stimulus;* each time he makes a 'wrong' response, he learns something else, *but he does not learn anything about the 'correctness' or 'wrongness' of the to-be-finally-learned set of discriminanda.* Eventually he gives up responding to his first set of discriminanda and responds to another set, and another set, etc., until he begins to respond to the relevant set. From then on, and from then only, is he learning anything about the discrimination involved, or, from then on only are his 'bonds' being strengthened, etc." (**9**, p. 111).

As an alternative to this interpretation, the writer has proposed a theory based on conditioning or association principles (**15**, **16**), which Krechevsky (**9**) has included among the group which he has named the continuity theories in contrast to the non-continuity formulation. According to the writer's version of this type of continuity theory, discrimination learning is conceived of as a cumulative process of building up the excitatory tendency or association between the positive stimulus cue and the response of approaching it, as compared with the excitatory tendency of the negative stimulus cue, which receives only non-reinforcement, to evoke the response of approaching it. This process continues until the difference between the excitatory strengths of the two stimulus cues is sufficiently great to overshadow always any differences in excitatory strength that may exist between other aspects of the situation which happen on a particular trial to be allied in their response evoking action with one or the other of the cue stimuli. For example, the difference in the excitatory strengths of the cue stimuli to their responses must, for learning to be complete, become greater than the difference between the excitatory strengths of the position cues, e.g., the left and right alleys or food boxes.

As a critical test of this continuity theory and an alternative insight theory [1] which was elaborated on the basis of Lashley's suggestion, the writer (**15**) proposed the following experiment. If, in a discrimination situation the significance of the stimulus cues is re-

[1] This insight hypothesis was essentially the same as that subsequently put forward by Krechevsky (**8**). Both merely are logical elaborations of Lashley's original notion.

versed, i.e., the positive stimulus made negative and vice versa, before the animal begins to show any learning whatever, the subsequent learning of the reversed problem would not, according to the non-continuity theory, be slowed up. The implication of the writer's continuity theory, on the other hand, was that the learning of the reversed group would be retarded when compared with a control group which did not have such initial reversed training.

The results of an already existent experiment on weight discrimination by McCulloch and Pratt (**14**) were then cited by the writer as supporting the continuity type of theory and as opposed to the non-continuity theory. These investigators found that the group in which the cues were reversed subsequently required a longer time to learn than the control group. This result seemed to represent a clear cut refutation of the non-continuity type of hypothesis. But then the supporters of this viewpoint came to its defense with a variety of devices including new experimental data and restatements of the issue which completely departed from the original one (**2**, **9**).

First, Krechevsky (**8**) called the McCulloch and Pratt data into question on the basis that their *S*s might have been responding on the basis of two hypotheses, one of which was the weight cue. As McCulloch (**12**) was quick to point out, however, this was not possible for, with the exception of one case, their *S*s did not respond more than 50 percent of the time to the positive stimulus before reversal. Obviously, then, they could not have been responding systematically to either one of the weight stimuli in combination with any other cue.

Next Krechevsky (**9**) reported the results of a new experiment with visual discriminanda which were, in part, contrary to McCulloch and Pratt's findings. In a situation in which the stimulus cues were cards containing horizontal and vertical rows of black squares on a white background, an experimental group, in which the cues were reversed for 20 trials, showed no difference in speed of subsequent learning from the control group. A second experimental group given 40 trials before reversal, however, did show retardation in the subsesequent learning of the reversal problem.[2] Accepting the results of the first experimental group as supporting the non-continuity theory, Krechevsky attempted to explain away the embarrassing results of the 40-trial group on the basis of the assumption that the pre-solution period occupied less than 40 trials. Just why this assumption was made is difficult to understand. The different implications of the two opposing theories were originally based on the assumption that the reversal occurs "before the animal begins to show any learning *whatever* . . ." (**15**, p. 444). All one had to do to verify the meeting

[2] The differences were significant at the one or five percent levels of confidence in the case of the three error scores, but were not significant for the trial score.

of this condition was to ascertain whether the *S*s had begun to respond more than chance to the positive stimulus. As both McCulloch (**12**) and Spence (**18**) have pointed out, however, Krechevsky presented no evidence that his *S*s were doing so. Moreover, McCulloch showed that the control group gave no evidence of responding above chance (50 percent) to the positive stimulus during trials 21–40, let alone responding systematically to it. It should be clear that the theoretical issue cannot be settled by such evasions. There is only one experimentally testable definition or criterion of the concept 'pre-solution period,' namely the occurrence of the correct systematic response tendency. The fact that the four-day group did not exhibit such a response tendency necessarily means that they were still in the pre-solution period. The results of this part of the experiment are thus seen to favor the continuity theory.

Turning now to the results of the 20-trial group, the writer has elsewhere (**18**) dealt at length with the fact that in visual situations involving form discrimination a new factor is introduced which is not present in the weight discrimination situation or in the case of brightness discrimination with the modern types of apparatus. Thus, when pulling in the differently weighted boxes, McCulloch and Pratt's *S*s inevitably received discriminably different weight stimulations. Likewise, in brightness discrimination it is practically impossible for the *S* to respond and not receive discriminably different retinal stimulations from the two stimulus sources (assuming, of course, we have differences which are supra-threshold). In such an experiment as Krechevsky's, however, it is quite conceivable that his *S*s did not receive discriminably different (proximal) visual stimulations from the two stimulus cards. Fixating the bottom portion of the stimulus cards, as the rat does at first in the Lashley jumping apparatus, the two stimulus patterns produced on its retina were very likely not discriminable. So long as the *S* continued this fixation, the retinal patterns of stimulation received from the two cards would not be distinguishable, and no differential learning with respect to the two cards would take place. In order for discrimination to develop in such a situation, it is necessary for the animal to learn first to make the appropriate receptor-orienting acts that will lead to the reception of the critical stimulus patterns or, to state it another way, that will lead to the reception of retinal stimulus patterns that are discriminable by the rat. Apparently in the Krechevsky experiment the *S*s learned to make the appropriate receptor-orienting response sometime after the 20th trial and before the 40th.[3]

[3] Two previous attempts have been made by the writer to clarify this problem of the reception of the relevant stimulus cue by the *S* (**16**, **18**). However, it is apparent, from such recent comments on the issue as those of Leeper (**10**) that this has been unsuccessful. Possibly the fault

It is obvious from the above discussion that our inability to specify the proximal retinal stimulus pattern in the case of form discrimination makes it impossible to employ this type of visual discrimination situation to settle the present theoretical issue. A more thorough analytical approach to the problem would have precluded its use in the first place. There is, nevertheless, a perfectly feasible way of testing the theoretical issue in a visual discrimination situation. It consists in the use of brightness cues rather than figure or pattern discrimination, for in the brightness problem there is no difficulty in arranging experimental conditions that will insure reception of the differential stimulus cues from the beginning of practice. Moreover, such an experiment provides a test of the theoretical issue in a learning situation in which the perceptual behavior is as simple as possible.

An investigation using brightness discrimination was recently carried out by Bollinger (1) as a master's thesis under the writer's direction. This experiment compared the learning of a white (+) versus black (−) discrimination problem by albino rats following two different pre-learning conditions. The experimental group of *Ss* was trained during this period for 30 trials on the reverse problem (black positive), while the control group was run for 30 trials with 50 percent reinforcement on black and 50 percent reinforcement on white.[4] Only *Ss* that exhibited chance responses during this pre-learning period were run in the subsequent learning problem. The results of this experiment were in complete agreement with those of McCulloch and Pratt. Both errors and trials were greater for the experimental group than the control group. The differences were significant beyond the one percent level as shown by the values of the *t* statistic.

Bollinger also ran a second experiment in which the experimental group was given only 20 trials in the pre-reversal training. Again the results were contrary to the implication of the non-continuity theory, for the experimental group required a significantly greater number of trials and errors than its control group. The *t* statistic for errors was again significant at the one percent level while that for trials was significant at the five percent level.

II. Statement of the Present Problem

In the experiment of Bollinger it was found that not all *Ss* exhibited a definite systematic response tendency or hypothesis during

has been the writer's use of such expressions as "not seeing the stimulus cue," or "the animal learns 'to look at' one aspect of the situation." If this surmise is correct, it is all the more surprising in view of the extent to which the non-continuity theorists favor such anthropormorphic descriptions.

[4] The purpose of the preliminary training with the control group will be discussed in the theoretical interpretation of the present experiment.

the 20 (30) trial pre-learning period. While no differences appeared in the subsequent learning of those *S*s that did exhibit such hypotheses and those that did not, it was decided to conduct a further investigation to ascertain whether, in a brightness discrimination situation, any association would be set up between the color cue and the reinforced response to it during a period in which the *S* was exhibiting a definite hypothesis with respect to some other cue, e.g., a dominant reaction to position. In other words, if an *S* has a systematic position preference, does consistent differential training to the color cues (e.g., reinforcement on black, non-reinforcement on white) have any subsequent effect so far as learning to respond to these cues is concerned? Examination of the formulation of Krechevsky quoted above will be seen (1) to emphasize the presence of such hypotheses and (2) the fact that the *S* does not, while so responding, learn "anything about the 'correctness' or 'wrongness' of the to-be-finally-learned set of discriminada" (9, p. 111).

The problem was investigated by first establishing a position habit in the *S*s, and then for a 20-trial period reinforcing one group of *S*s (experimental) consistently on one of the color cues, while a second group (control) was not reinforced differentially, but only 50 percent of the time on *each* of the color cues. The presence of a difference in the subsequent learning of the brightness discrimination by the experimental and control groups would suggest that an association was developed between the response and the color cue during the dominance of reaction to the spatial position. Absence of such a difference would be an indication that a spatial 'hypothesis' prevents the association of the simultaneously acting color cues with the response. The latter finding would lend support to the non-continuity type of theory, the former result to the continuity theory, as the writer understands these theories.

III. Experimental Procedure

A. Subjects

In all, 44 albino rats, bred in the laboratory of the psychology department of the University of Iowa, were employed. The ages of the *S*s at the beginning of the experiment ranged from 75 to 90 days. Half of the *S*s were males and half females. They were distributed equally between the experimental and control groups. Four *S*s were discarded from the experiment because of brightness preferences as will be described later in the discussion of results.

B. Apparatus

The apparatus differed somewhat from previous types employed with white rats. It was designed (1) to eliminate the necessity of handling the rats except at the beginning of each day's trials, (2) to insure the reception of the stimulus cues by the *S*, and (3) to give immediate reward or non-reinforcement upon response.

The principal parts of the apparatus are shown in the ground plan of Fig. 1. A turntable

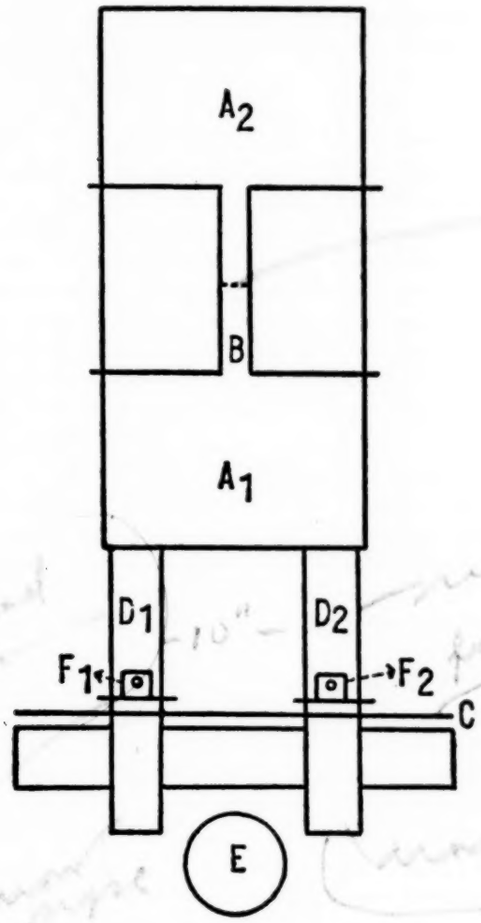


FIG. 1. Diagram of the apparatus.

divided into two portions, A_1 and A_2, by means of partitions, stood four feet above the floor and moved on a pivot made by two iron pipes. The smaller pipe, which was fastened to the table, fitted inside a larger one, which was based on the floor. The fit of the pipes was sufficiently close so that there was practically no play in the table, and yet it could be easily rotated by means of a stock attached to one of its sides. A one-foot alley (B), two in. in width, connected the two open sides of the table. A vertically sliding door in the middle of this alley permitted the E to allow the S to pass from one side of the apparatus to the other.

The second main part of the apparatus consisted of a stand which provided a one-way screen (C) behind which E was situated and which supported two elevated runways (D_1, D_2) $3\frac{1}{2} \times 10$ in. The latter were interchangeable; one was painted white and the other black. The food box (F_1, F_2) consisted of a hole, $\frac{3}{4}$ in. in diameter, drilled in a block of wood ($2 \times 3\frac{1}{2} \times 4$) placed at the end of the elevated runway. Immediately behind this block a vertical piece of plywood ($6'' \times 6''$) was fastened. The runways were fitted into grooves which permitted them to be moved back a few inches and thus permit the rotation of the apparatus between trials. A distance of 10 in. separated the two runways.

In addition to the white and black runways, two duplicate runways painted an intermediate grey were provided. These were used in the preliminary training and in certain other parts of the experiment as will be described in the next section.

C. Procedure

The experiment may be divided into five periods: (1) preliminary training, (2) position habit training, (3) pre-solution training, (4) position reversal training, and (5) the learning problem.

Before the preliminary training was begun the subjects were placed on the turntable in groups of two to four in order to familiarize them with it. The *preliminary training* involved

two days. On day 1 with the grey runways in place each *S* was placed on the left (or right) runway and permitted to eat food. On day 2 the animal was placed on the side of the platform away from the food runways and trained to run through alley *B* to the food boxes upon the raising of the door.

Each *S* was next given three days of *position habit training*. Employing the grey runways, 10 trials were given each day with food being available on only one side. Half of the *S*s were trained to the left and half to the right position.

The *pre-solution training* employed the white and black runways. It differed in procedure for the experimental and control group *S*s. The former were reinforced (i.e., received food) only when they chose the black pathway (half of the group were reinforced only on white). This meant that an *S* which persisted in its position habit was reinforced 50 percent of the time as the white and black runways were presented equally often on the left and right side. Twenty trials in all (two days) were given during this period. The control-group *S*s were not reinforced consistently on either the white or black pathways during this period, but instead were reinforced 50 percent of the time on black and 50 percent on the white. As in the case of the experimental group, the white and black runways appeared equally often in each spatial position. A prearranged position of the white and black cues was followed in which each was reinforced equally often on each side in the 20 trials. In all periods of training no correction of an erroneous response was ever permitted. Instead the turntable was rotated and after the usual period of time between trials (approximately 30 sec.) the next trial was given.

Position reversal was carried out on the following day. With the grey runways present, food was placed on the side opposite to that on which the original position habit was established. Training was continued until the *S* chose the newly rewarded position for three successive trials, or for four out of five trials.

In the *learning period* both groups of *S*s learned to respond to one of the color cues, white or black. In the case of the control group half of the *S*s were trained to white and half to black. In the case of the experimental group, those rewarded on the black runway during the pre-solution period were trained on the white path during the learning period, while those rewarded on the white during the pre-solution period were required subsequently to learn to go to the black. Ten trials were given daily and training was continued until the criterion of 90 percent correct in 20 trials was attained.

IV. Results

The development of the position habit in both groups of *S*s, experimental and control, during the initial three day period is shown by the data in Table I. It will be seen that the mean strength of the preference was approximately the same for the two groups during the first three days. Days 4 and 5 in this table represent the position responses of the pre-solution period. Although both groups of *S*s continued to respond to their preferred position during the 20 trials of this period, there was a noticeable decrease in the percentage of such responses, particularly in the case of the experimental group. The

TABLE I

Showing Development of Position Responses

Day	Percent Responses to Preferred Side	
	Experimental Group	Control Group
1	86.0	85
2	91.5	90
3	94.5	94.5
4	78.0	82.0
5	78.0	86.0

relation of the maintenance of the position hypothesis during this period to the subsequent learning of the discrimination problem will be discussed later.

During the pre-solution period none of the Ss responded more than 65 percent of the trials to either the black or white stimulus. As was indicated in the description of the Ss, four animals were discarded from the experiment at this point because they exhibited a preference greater than this amount for either the black or white stimulus. Two of these were trained under the control condition in the pre-solution period and two under the experimental condition.

The data on the learning of the discrimination problem in terms of errors and trials are presented in Table II. It will be observed in

TABLE II

DATA ON LEARNING OF DISCRIMINATION PROBLEM

Group	Errors		Trials	
	Mean	σM	Mean	σM
Experimental	41.00	3.11	95.00	5.60
Control	26.05	2.55	70.50	4.66
Difference	14.95		24.50	
t	3.72		3.36	

the case of both measures that contrary to the non-continuity theory the experimental group was considerably retarded in its learning when compared with the control group. That the differences are statistically significant is shown by the *t*-values of 3.72 obtained for errors and 3.36 for trials. A *t*-value of approximately 3.5 is at the one tenth of one percent level of significance with this number of degrees of freedom.

It should be noted that the *t*-test is in this case a conservative one. Because of the fact that simple random samples were not employed, but instead the two groups, experimental and control, were equalized in number for the sex and direction of training variables, there is no possibility of a systematic effect of these variables on the differences in the means of the experimental and control groups. However, the error estimate allows chance variations in these factors and hence is inflated. To the extent, then, to which these two controlled variables of the present experiment are related to speed of discrimination learning, the difference obtained is even less likely to be due to chance factors than is indicated by the above *t*-analysis.

A more appropriate statistical analysis of the results of the present experiment, which was of the factorial type, is provided by the method of analysis of variance. Such a statistical analysis provides a better

indication of the significance of the difference between the two groups because it permits the elimination of any sex and direction of training differences from the error variance. Analysis of variance also permits one to test for any interaction of the three variables.

When all possible interactions were tested, none were found to be significant. This means that the *differences* between the experimental and control groups did not differ significantly for the two sexes or for the direction of training, i.e., the differences were essentially the same for both males and females, etc. Using the within groups variance as the error estimate, it was found that only one *F*-ratio, that for the principal experimental factor, the pre-solution methods (experimental versus control procedure) proved to be significant. In the case of errors, this *F*-ratio was 14.37, and for trials it was 12.28, which for the appropriate degrees of freedom is significant well beyond the one percent level of confidence (.01 = 7.50). It may be concluded from this that the differences between the mean number of errors and trials for the experimental and control groups cannot reasonably be attributed to chance fluctuations in sampling.

Fig. 2 presents a further comparison of the learning of the experimental and control groups. This graph shows the percent of correct responses in the learning of the discrimination problem *by the two sub-groups which had the white stimulus positive.* One of the most interesting aspects of this graph is the effect of the elimination of the position habit on the experimental group. During the pre-learning period while these *S*s were responding predominantly to a position cue (approximately 80 percent of trials) they chose the white stimulus 51 and 56 percent in the successive 10-trial blocks. They were, of course, reinforced during this period every time they responded to the black card. Following the elimination of the position preference, these same *S*s responded correctly (i.e., to the white stimulus) 37, 46, 46, and 42 percent of the time, on the first four successive 10-trial periods. This result would suggest that during the pre-solution period there was some kind of learning or association formed with respect to the black stimulus, and then when the masking position preference was eliminated it was able to exhibit itself.[5] Comparison of the two curves of learning suggests, furthermore, that the experimental group takes longer because it starts from a lower value. From the 50 percent point on, the two curves pretty well parallel each other.

One final point with respect to the experimental data. Examination of the responses made during the pre-solution period with the

[5] If the hypothesis is made that the mean number of responses to the positive stimulus by the experimental group in the first 10 trials of the learning problem is 5.0 (or 50 percent correct), the obtained mean of 3.7 (σ_m = .47) departs 1.3 responses and provides a *t*-value of 2.76. With a sample of this size (d.f. = 9) the chances would be less than 5 in 100 of getting a *t* as large as 2.76 by random fluctuations in sampling.

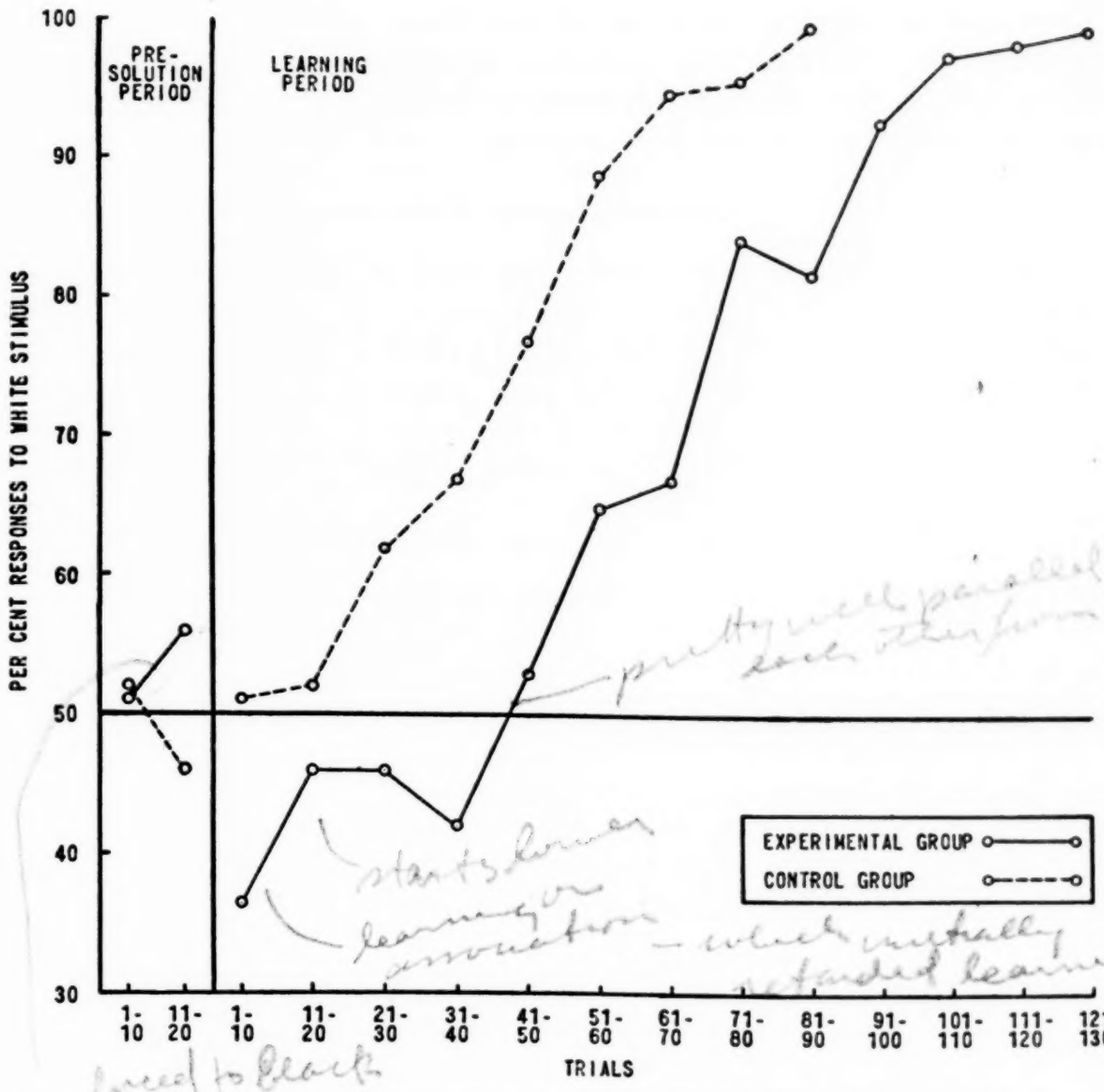


FIG. 2. Pre-learning and learning data for *Ss* trained to the white stimulus.

brightness stimuli revealed the fact that not all *Ss* maintained the position habit throughout the 20 trials. The question immediately arises as to whether there is any difference in the learning of the subsequent discrimination problem between those *Ss* of the experimental group that did maintain their position hypothesis and those that did not. Analysis of the data revealed that 10 of the 20 *Ss* maintained a position preference, averaging (as a group) 92 percent of their responses to the preferred side in their last five trials. These 10 *Ss* averaged 41.4 errors and 98 trials, which values, it will be seen by comparison with Table I, differ little from the means for all experimental *Ss*. Furthermore, it was found that seven *Ss* of this experimental group that did not maintain their position habit throughout the 20 trials (in fact, they responded only 52 percent of the last five trials to the favored side) made about the same learning record in the subsequent discrimination problem. These seven *Ss* had a mean of 43

errors and required an average of 100 trials to learn. It should be obvious from this that the presence or absence of a position habit during this period made no difference whatsoever in the learning of the subsequent discrimination problem.

V. Discussion of Results

The issue between the continuity and non-continuity theories, stated in the form of a question, is as follows. Does the animal before it learns a discrimination problem, i.e., before it responds systematically to the positive stimulus cue and while it may be responding systematically to other stimulus cues, learn anything about (form any associations with) the positive stimulus cue? The non-continuity theory, building on the suggestion of Lashley, answers the question in the negative. During this pre-solution period, it holds, the animal may be responding systematically to other cues, e.g., position, and thus learning about them, but it does not learn anything about the 'to-be-finally-learned set of discriminanda.' Not until the *S* begins to respond to the relevant set of discriminanda, says Krechevsky, does it learn anything about the discrimination involved or "from then on only are his 'bonds' being strengthened, etc." (**9**, p. 112). On the other hand, the continuity formulation, holding that the formation of associations is a continuous process, would insist that so long as the animal received discriminably different stimulation from the stimulus objects, differential associative tendencies would be developed with respect to them.

Consider, now, the results of the present experiment. The question of whether any associations were formed between the positive stimulus of the pre-learning period in the case of the experimental group can be answered in two ways: first, by the occurrence of any tendency to prefer this stimulus following the removal of the masking position habit and, secondly, by the slower learning of the discrimination problem on the part of the experimental group as compared with the control group. Both of these experimental implications were born out by the present data as Table I and Fig. 2 show. This evidence points to the interpretation that in the experimental group the consistent reinforcement of one of the stimulus cues had a definite effect. Contrary to the suggestion of Lashley, associations apparently were being formed with the relevant stimulus cue even though the *S*s were responding on the basis of some other hypothesis.

Attention must be called now to another unfortunate misunderstanding of this experimental issue that has characterized the writings of the proponents of the non-continuity theory (**9**, **11**). In comparing his two experimental groups with his control group, Krechevsky presented his experimental data in two different ways. First, he com-

pared the learning of his experimental groups after reversal with the total learning record of the control group and, secondly, he compared the data of the experimental groups after reversal with the learning of the control group *after* 20 (or 40) trials. The present writer has attempted to point out the inappropriateness of the former comparison (**18**, p. 279). If our interest is in ascertaining whether the learning of the experimental group in the subsequent discrimination problem is retarded *because of the reversed conditions of reinforcement (with its consequent hypothetical formation of conflicting associations) during the pre-learning period*, then it is not appropriate to compare the learning scores of this group with the total learning scores of the control (non-reversed) group. The comparison must be made with the control group scores *after* 20 (or 40) trials. The reason for this should be obvious.

In order to answer the question asked we must attempt to hold all variables constant but the single factor we are interested in. Including the data of the first 20 trials of the control group in the measures to be compared would fail to do this, for it would mean that we would be comparing the learning record of a group (experimental) *after* it had 20 (or 40) trials in a discrimination situation with a group (control) which did not start with this advantage. This advantage would tend to produce lower learning scores for the experimental group than for the control group and thus tend to cover up any retardation in the learning of the experimental group that might result from the development of conflicting associations.

Thus, in the present experiment both groups of *S*s had an equal amount of experience in the discrimination situation prior to the beginning of the learning problem. Any difference in the rate with which the discrimination was learned must, then, be ascribed to the effects of the different operations employed in the pre-solution training. But according to the non-continuity theory the consistent reinforcement of response to the negative stimulus during this period should have had no such retarding effect. The results of this experiment, like those of Bollinger, are thus seen to present a clear-cut refutation of the non-continuity theory.

VI. Summary

Two groups of *S*s, 20 in each group, learned a white-black discrimination problem. Prior to the learning, both groups of *S*s were trained for 30 trials to their preferred side, the black and white cues being absent. The experimental group was then given 20 trials in which response to the negative stimulus of the learning situation was rewarded. The control group, on the other hand, was reinforced 50 percent of the time on both the white and black stimuli during this

20-trial pre-learning period. Following a further period of training in which the position preference was eliminated, the regular learning problem was presented. Contrary to the implications of the non-continuity interpretation of discrimination learning, comparison of the experimental and control groups revealed a significant difference in speed of learning This difference was interpreted as refuting the assumption of this type of theory that the animal learns nothing with respect to the relevant stimulus cue before it begins to respond systematically to it and while it is exhibiting some other 'hypothesis.' On the other hand, the results of the investigation are considered to be in agreement with the continuity type of theory, which postulates the learning to consist in the cumulative development of an association between the cue stimulus and the response, regardless of the presence of systematic response tendencies to other aspects of the stimulus situation.

(Manuscript received January 2, 1945)

References

1. Bollinger, P. An experimental study of the pre-solution period of discrimination learning in the white rat. M. A. thesis deposited in State University of Iowa Library, 1940.
2. Haire, M. A note concerning McCulloch's discussion of discrimination habits. *Psychol. Rev.*, 1939, **46**, 298–303.
3. Hamilton, G. V. A study of trial and error reactions in mammals. *J. anim. Behav. Monogr.*, 1911, **I**, 33–36.
4. Hull, C. L. *Principles of behavior.* New York: Hamilton Century, 1943.
5. Kellogg, W. N., & Wolfe, I. S. 'Hypotheses' and 'random activity' during the conditioning of dogs. *J. exp. Psychol.*, 1940, **26**, 588–601.
6. Krechevsky, I. 'Hypotheses' versus 'chance' in the pre-solution period in sensory discrimination-learning. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 27–44.
7. Krechevsky, I. 'Hypotheses' in rats. *Psychol. Rev.*, 1932, **39**, 516–532.
8. Krechevsky, I. A note concerning 'The nature of discrimination learning in animals.' *Psychol. Rev.*, 1937, **44**, 97–103.
9. Krechevsky, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, **45**, 107–133.
10. Lashley, K. S. *Brain mechanisms and intelligence.* Chicago: University of Chicago Press, 1929.
11. Leeper, Robert. Dr. Hull's *Principles of behavior.* *J. gen. Psychol.*, 1944, **65**, 3–52.
12. McCulloch, T. L. Comment on the formation of discrimination habits. *Psychol. Rev.*, 1939, **46**, 75–85.
13. McCulloch, T. L. Reply to a note on discrimination habits. *Psychol. Rev.*, 1939, **46**, 304–307.
14. McCulloch, T. L., & Pratt, J. C. A study of the pre-solution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18**, 271–290.
15. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
16. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1927, **44**, 430–444.
17. Spence, K. W. Solution of multiple choice problems by chimpanzees. *Comp. Psychol. Monogr.*, 1939, **15**, No. 3, Serial No. 75.
18. Spence, K. W. Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, **47**, 271–288.
19. Yerkes, R. M. The mental life of monkeys and apes: a study of ideational behavior in man and other animals. *Proc. nat. Acad. Sci.*, 1916, **2**, 631–633.

THE EFFECT OF ELECTROSHOCK CONVULSIONS ON THE MAZE HABIT IN THE WHITE RAT

BY CARL P. DUNCAN*

Brown University

I. Introduction

Since the introduction of convulsive therapy into psychiatry and its wide use in the treatment of psychogenic disorders, there have been several questions raised concerning the nature of the therapy and its subsequent effects. We have selected one of these questions for experimental study. This is: what is the effect of electroshock treatment on memory functions? This question might also be raised in connection with the other two major types of 'shock' treatment, the administration of metrazol and of insulin, but we shall restrict this study to the use of electricity.

Cerebral electroshock as treatment for the psychogenic behavior disorders was first introduced by Cerletti and Bini in 1938 (**9**). It has enjoyed a tremendous popularity, but as yet there have been few well-controlled experiments to determine the nature of treatment or any deleterious effects. One of the disadvantages of the therapy is the frequent report of patients that they are experiencing memory difficulties. Wortis (**16**) says that memory difficulties may persist for many weeks if 10 or more treatments are given. Smith, Hughes, Hastings, and Alpers (**14**), in a discussion of complications in electroshock treatment, report that almost all patients experience memory defect sometime during treatment and that the patients consider it an unpleasant symptom. These authors find that in some patients it is prominent; in several it has persisted for several months. They report no patients in whom the memory loss is permanent, and indicate that the loss is usually a spotty defect, ordinarily for proper names, places, and dates.

Few electrotherapists have experimented on their human patients for possible memory difficulties as a result of treatment. Levy, Serota, and Grinker (**6**) report a study on 21 of their own cases. Electroshock was used with 12 patients, while 11 received metrazol injections. They found that impairment in intellectual functions

* This paper is a portion of the author's M. A. thesis, which is on file in the John Hay Library, Brown University. The author is deeply indebted to Dr. Harold Schlosberg for help with this research.

occurred much more frequently in the electrically treated cases (8 out of 12). Impairment in electrically treated patients lasted from one to several weeks, and patients with more shocks showed greater defect.

Zubin and Barrera (**18**) taught paired word associates to several patients under various conditions. Some learned before taking the electric treatment, some during the course of treatment, and others after a series of shocks. Recall, recognition, and relearning were tested under the same conditions and thus control groups were obtained. They report that when no shock intervenes between learning and relearning, the number of trials saved was significant. However, when a shock is interpolated between learning and relearning, there is no saving, but rather a slight but nonsignificant loss. They emphasize that learning ability is no poorer after shock than before shock. For the recall tests they found that recall after shock was significantly less than the recall after the control and nonshock periods. With regard to recognition, the patients were able to recognize correctly both paired associates nearly 94 percent of the time under control conditions. Under shock conditions this percentage dropped to 81.5. These writers conclude with a report on recent and remote learning. Since some of the patients were taught material some time before treatment, and others immediately before treatment, a measure of the effect of the shocks on recent and remote learning was obtained. In this connection the authors say, "Apparently the shock affects the learning acquired immediately before shock more than the learning acquired less recently before shock" (**18**, p. 597).

We have found only one other study on human *S*s. Sherman, Mergener, and Levitin (**12**) first report that, with few exceptions, other investigators have found amnestic syndromes in their electrically treated patients with occasional severe loss. Sherman and her associates administered a series of six tests before, during, and after a course of electrotreatments in four patients. Their results showed that there was an increase in the average scores of the memory tests when the responses during the course of therapy were compared with those preceding the beginning of treatment. This increase, however, was insignificant except in one test—report of a paragraph heard—but the fact that there was a slight improvement in all tests suggests that the treatment had a favorable effect. They attribute the gain to an improvement in the previously impaired processes of attention and concentration.

Several experiments have reported the effect of electrical or chemical shocks on habits in animals. Kessler and Gellhorn (**4**) conditioned 18 rats to jump a partition, using an electric grid as the unconditioned stimulus. They then caused these responses to be ex-

tinguished, and when this was accomplished, some rats were subjected to electrical convulsions and others to metrazol shocks. They report that both electricity and the drug caused the conditioned responses to reappear. From this they postulate that in clinical practice shock may eliminate inhibitions.

A short experiment is reported by Rosen and Gantt (**11**) on the effect of metrazol convulsions on conditioned reflexes in dogs. They conditioned two dogs to make a salivary response, and two others to make a motor response. The metrazol had a more pronounced disrupting effect on the motor CR.

Page (**8**) reports several studies on electrically induced convulsions in cats and rats. In one experiment he tried to condition the convulsion itself by sounding a bell just before applying the electricity. No conditioning was obtained. In another experiment he studied the effects of convulsion on general behavior and weight loss in rats. Ten animals were reared in one cage, and at the age of nine weeks five of them were given several shocks. There were easily noticeable differences in general activity; the normal rats were much more active. The shocked rats did not compete for food, and lost weight during the series of shocks. Two months after the end of the experiment the shocked rats were still smaller than their fellows. Both of these results have been corroborated in the present experiment.

After Siegel (**13**) had taught two groups of rats a modified form of the Graham-Gagné runway, one group was subjected to 10 electroshock convulsions; the other group had the electrodes clamped to the ears once each day but received no shock over the same period. In the first retention trial both groups showed a slight loss of the habit but the loss was equal for both. In several subsequent relearning trials there was very little difference between the two groups.

Bunch and Mueller (**1**) have indicated that rats subjected to a series of metrazol convulsions manifest no differences in the subsequent learning of a 14-unit multiple-T maze.

Using an automatic multiple-T maze, Heron and Carlson (**3**) found an insignificant difference in their convulsion (metrazol) group.

Loken (**7**), however, has reported fairly large time and error increases in his animals' relearning of a six unit maze. He also used metrazol.

Finally, Stainbrook and Löwenbach (**15**) have observed no disturbance in the retention of a simple water maze habit in rats subjected to repeated electroshocks.

Evidently the possibility of malfunctioning of the memory process, as a result of electricity through the head, is not to be ruled out, even though there is not unanimous agreement. We shall show in this

experiment results which lead us to believe that electroshock convulsions produce at least one form of memory loss.

II. Experimental Procedure

The maze used was a reproduction of Lashley's maze III described in his *Brain mechanisms and intelligence* (5). Many of the animal experiments which we have cited employed a relatively simple habit, as, for example, Siegel's use of the Graham-Gagné runway. This may be the explanation of the negative results often obtained. We therefore decided to vary procedure by using a more difficult habit, and chose this particular maze because Lashley had found it difficult for his rats. The maze is made up of four parallel alleys with the dead ends of each alley forming the culs, making 8 culs in all. The rat must completely reverse his direction as he goes from one alley to the next to avoid a cul.

The animals were 28 male albino rats 170 to 180 days old at the beginning of the experiment.

The rats were first trained to run an 8½ foot alley runway, one trial per day for 15 days. They were fed wet Purina mash for 10 min. after the daily trial. This was the reward and daily ration for all animals. All animals were then started on the maze, running one trial a day with food as reward. As each rat learned the maze to a criterion of three errorless trials in succession he was started on the electroshock procedure. Time, trial, and error scores were kept.

The shock apparatus consisted of a voltage divider connected to the 110 volt A. C. lighting circuit. The divider was set at 85 volts, the 85 volt point being determined with a voltmeter and repeatedly checked over a period of days. Wires led from the divider to 'alligator clamp' electrodes which were wound with gauze. Before each rat was shocked the electrodes were thoroughly soaked in physiological salt solution. The time of application of the current was approximately .22 sec. The duration was determined by the time it took the experimenter to flick a toggle switch on and off quite rapidly. After some practice there was little deviation in the time. Stopclock measurements showed it to vary between .20 and .25 sec.

The animals were separated into three groups as they reached the criterion. The Experimental Shock Group, Group E, received one shock through the head each day for 30 days. In this group the electrodes were attached to the ears. It is these animals that we expect to exhibit memory loss for the habit. Group S was the Shock Control Group. These rats received the same voltage for the same period but the electrodes were attached to the back legs. If the electricity does not have a specific effect on nerve tissue, we would expect the same decrement to memory no matter where the electrodes were placed on the body. Further, the post-shock behavior of rats shocked through the brain is much different from that of rats shocked through the legs, as we shall indicate later. We therefore introduced Group S to control some of these differences in post-shock behavior. Group R was the Retention Control Group. This last group received no shocks for the 30-day period. Food was always in the living cages of all animals during the electroshock period.

Since we feared that shock through the brain might kill some of the animals, 12 rats were placed in Group E. Groups S and R each consisted of eight rats. The groups were equated as nearly as possible in terms of learning ability on the maze.

At the completion of the 30-day period for each rat he was again started in the maze and kept running it until he had relearned it to the original criterion of three errorless trials. The same procedure was used in relearning as in learning.

III. Results

The results are presented in Tables I through VI and in Figs. 1 and 2. In Table I the mean, the median, and the standard deviation for errors, trials, and time expressed in logarithms [1] are presented for the three groups of animals. Data are shown for both learning and relearning.

[1] The argument for converting certain skewed scores such as time into logarithms is presented in a paper by Graham and Gagné (2).

The mean scores of the three measures of learning (from Table I) were tested for significance of differences by employing Fisher's test. These *t*-scores with their accompanying probability values (*P*) are given in Table II.

TABLE I

RATE OF LEARNING AND RELEARNING

Learning

Group	Errors			Trials			Log Time		
	Mdn.	M.	SD_m	Mdn.	M.	SD_m	Mdn.	M.	SD_m
(E) Exp. Sh.....	23	32.4	18.4	11.5	12.6	6.7	2.53	2.58	2.42
(S) Sh. Con.....	22	24.6	9.6	8	10	4.2	2.20	2.29	1.98
(R) Ret. Con.....	26	29	14.8	9	10.2	4.9	2.26	2.35	2.12

Relearning

Group	Errors			Trials			Log Time		
	Mdn.	M.	SD_m	Mdn.	M.	SD_m	Mdn.	M.	SD_m
(E) Exp. Sh.....	4	9.8	16.6	7	8.8	6.1	2.69	2.72	2.55
(S) Sh. Con.....	2	3.3	2.7	2	1.9	1.3	1.83	1.97	1.85
(R) Ret. Con.....	1.5	1.4	0.09	1	1.1	0.9	1.38	1.53	1.43

TABLE II

DIFFERENCES BETWEEN GROUPS FROM DATA OF TABLE I

Learning

Groups	Errors		Trials		Log Time	
	t	*P*	*t*	*P*	*t*	*P*
E and S.....	.96	below 50%	.85	below 50%	1.29	above 20%
E and R.....	.39	above 50%	.79	below 50%	1.22	above 20%
S and R.....	.44	above 50%	.08	above 50%	.78	below 50%

Relearning

Groups	Errors		Trials		Log Time	
	t	*P*	*t*	*P*	*t*	*P*
E and S.....	3.06	below 1%	3.36	below 1%	4.53	below .1%
E and R.....	3.46	below 1%	3.27	below 1%	9.76	below .1%
S and R.....	1.40	below 20%	1.10	above 20%	2.40	above 2%

The percentages of errors, trials, and time saved in relearning were computed for each rat, and the mean percentages saved are presented in Table III. The *t*-test was also applied to the mean percentages of Table III, and the resulting data with the corresponding *P* scores are given in Table IV.

TABLE III

SAVINGS OF RELEARNING OVER LEARNING EXPRESSED IN MEAN PERCENT

Group	Errors	Trials	Log Time
(E) Exp. Sh.	18.2%	22.8%	−11.1%
(S) Sh. Con.	84.4%	80.5%	16.1%
(R) Ret. Con.	94.1%	84.5%	36.5%

Fig. 1 represents learning and relearning curves in terms of errors per trial for a typical rat from each of the three groups. Fig. 2 shows similar curves in terms of log time per trial. The rat chosen to represent each group is that nearest the median for both learning and relearning. The graphs are presented solely to illustrate the course of learning and relearning; our conclusions will rest on the tabulated material. For such illustrative purposes we believe that the learning curves of actual animals are more satisfactory than are average curves.

TABLE IV

DIFFERENCES BETWEEN GROUPS FROM DATA OF TABLE III

Group	Errors		Trials		Log Time	
	t	*P*	*t*	*P*	*t*	*P*
E and S	3.7	below 1%	3.2	below 1%	3.4	below 1%
E and R	4.7	below .1%	3.5	below 1%	7.3	below .1%
S and R	1.5	below 20%	.42	above 50%	2.5	below 5%

We shall not analyze these data in detail. The results shown in Table II indicate that during original learning the groups were fairly well matched in performance. For errors the probability is about .50 that the groups are unequal. For trials the probability is about the same as for errors, approximately .50. For time the differences among the means of the groups are greater, but since the probability is still no greater than .20 we can safely assume that the groups did not differ markedly in original learning, before shocking.

After the animals had been shocked, however, differences showed clearly in relearning scores. The Experimental Shock Group of rats, after cerebral electroshock, showed a very significant inferiority to both of the other groups. In terms of both errors and trials during relearning the differences are significant beyond the one percent level; in time they are beyond the 0.1 percent level. On the other hand, the Shock Control Group, which had received the same charge through the legs, was only slightly inferior to the Retention Control Group that had received no shock. In terms of both errors and trials the differences between the two control groups are significant at only

the 20 percent level. The time difference is fairly significant, approaching the two percent level.

We can safely conclude that the shocks through the head considerably prolonged relearning, showing in terms of errors, trials, and time. Further, similar shocks through the legs may have had some detrimental effect on relearning, but the smallness and variability of the differences do not indicate this definitely.

The percentage-saving data presented in Tables III and IV show these differences slightly better than do the raw score data. For

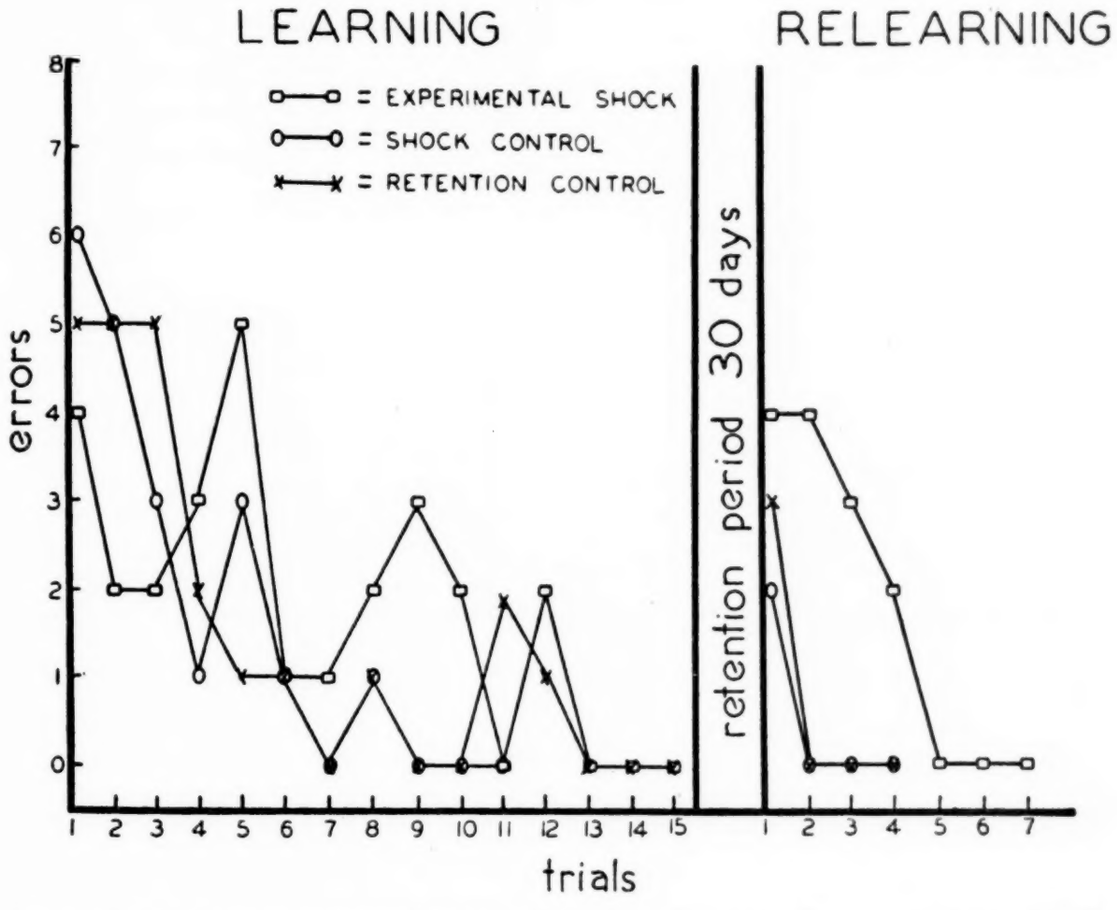


FIG. 1. Errors made on each trial by the median rat of each of the three groups during learning and relearning.

errors, Group E has a mean savings of 18.2 percent as compared with the mean savings of Group S of 84.4 percent, and the savings of Group R of 94.1 percent. That these differences are very significant is indicated in Table IV. Between the mean savings of Groups E and S, the *P* value is below one percent, and between E and R it is below 0.1 percent. Between the control groups, S and R, the difference is not significant, being at the 20 percent level.

In terms of trials the mean percentages saved are about the same as for errors, and the *P* of .01 between Group E and the control groups

again indicates the great effect of shocks through the head on relearning. The control group difference is not significant.

The percents saved for time are consistently smaller than those for trials and errors. Group E actually had a loss of − 11.1 percent, while Group S saved only 16.1 percent, and Group R saved 36.5 percent. As can be seen from Table IV, however, these differences are also clearly significant. Between E and S there is a P of .01 and between E and R a P of 0.001.

Thus, when the data are computed in percent saved, Group E was significantly worse for having been shocked through the head, and again there was a slight but inconsistent decrement in Group S because of leg shock. The savings data are particularly important,

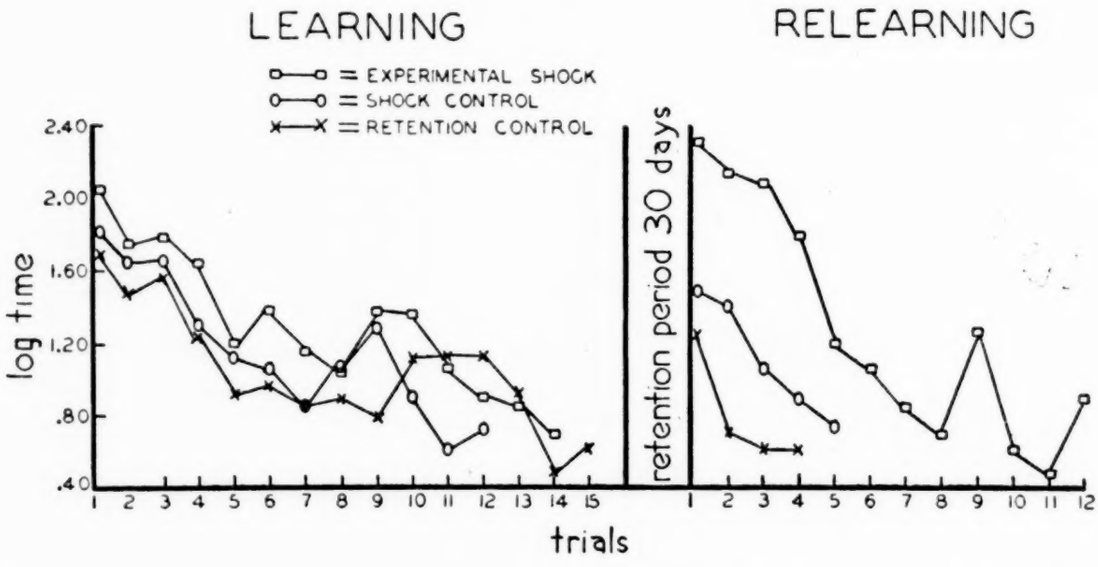


FIG. 2. Log time for each trial for the median rat of each of the three groups during learning and relearning.

since they are independent of any possible mis-matching of original ability among the groups.

Fig. 1, in which errors are plotted as a function of trials, shows that the median rats of Groups E and R required 15 trials to reach the criterion of three errorless trials, while the Group S rat required 11 trials to learn. The graph shows the great amount of overlap among the three groups during learning. When the scores of the same rats are plotted for relearning, however, both of the control rats were errorless on the second trial and both reached the criterion again on the fourth trial. The experimental rat was not errorless until the fifth trial and attained the criterion on the seventh trial.

Fig. 2 plots log time against trials for the median rat for each group during learning and relearning. Here also there is overlap in the times of the three rats during learning, indicating that the groups

are comparable. In relearning, however, the Group E rat took several more trials to reach criterion and his time for each trial is long when compared with the control rats on the same trial. The clear difference between the two control groups in time during relearning, as suggested in Fig. 2, is, according to Table II, significant above the two percent level. This was the greatest difference found between Groups S and R.

The results presented so far have all been in terms of relearning scores. These data do not give us an unambiguous measure of retention since the electroshock might affect only relearning ability. We have therefore presented in Table V the mean errors and mean log time for the groups on the first relearning trial, as a more direct measure of retention. The differences between these means are shown in

TABLE V

DIFFERENCES BETWEEN GROUPS IN RETENTION AS MEASURED IN THE FIRST RELEARNING TRIAL

	Mean Errors	Mean Log Time
Group E	4.90	2.12
Group S	1.71	1.44
Group R	.875	1.04

TABLE VI

SIGNIFICANCE OF DIFFERENCES BETWEEN GROUPS FROM DATA ON TABLE V

	Errors		Log Time	
	t	*P*	*t*	*P*
E and S	4.41	below .1%	4.30	below .1%
E and R	5.99	below .1%	6.42	below .1%
S and R	1.43	below 20%	2.38	below 5%

Table VI. From Table VI we see that the differences between Group E and the controls are significant at the 0.1 percent level for both errors and time. The retention differences between Groups S and R are not significant, below the 20 percent level, for errors, and only fairly significant, below the five percent level, for time.

IV. Discussion

These results indicate clearly that the passage of electricity through the head and brain had a markedly detrimental effect on a recently learned difficult habit. On the other hand, the same dosage of electricity through the hind legs had but little effect. This is the opposite of what one would expect from the hypothesis that the therapeutic action of electroshock on mental patients is due to the 'psychic jolt' and the fear of punishment. Our experimental rats that were shocked through the skull exhibited almost exactly the

same behavior shown by psychotics who receive electrotreatment. There was immediate immobility upon passage of the current (in human terms, 'loss of consciousness'), followed, with a latency of a few seconds, by first a tonic and then a clonic convulsion of the whole body. This muscular discharge was followed by a slow return to the normal state where external stimuli produced a response, a process requiring several minutes. There was no fear reaction even after the animal had received most of its quota of 30 shocks.

In the Shock Control animals, however, the response to the shock through the legs was quite violent. Squealing, defecation, and micturition always accompanied the brief charge. After two or three shocks the rats showed fear when *E*'s gloved hand picked them up. This was not observed in the experimental rats. It would seem that animals getting shocked through the legs, showing fear, and retaining 'consciousness' during passage of the current would, according to the 'psychic shock' theory, lose recently established habit systems such as our maze. This did not occur to any great extent. The small loss shown by the Shock Control Group can probably be explained on the basis of the conditioned fear of *E*, since these rats were the only ones that showed any dislike of being handled and then only after several electroshocks. But the unfrightened electroshock rats were the ones that showed the large deficit.

There is still very little known about the underlying physiological processes that occur when an organism, and especially its brain, is subjected to electricity, insulin, or metrazol. Just as this experiment was completed, Riess and Berman (**10**) published a paper on the effect of insulin on habits. Some of their several experiments were quite similar to ours, except for their use of insulin. They hypothesize that the known effect of insulin on the brain, cerebral anoxia, tends to break up recently formed habits more than older ones. Their several experiments led to the following conclusions: 1. Insulin has a greater disintegrative effect on the learning of a less well fixed habit than on a habit of greater fixation. 2. Insulin has a greater detrimental effect on a longer and more difficult maze than on a shorter and easier maze. 3. Insulin has a disintegrating effect on the initial acquisition of a maze habit when compared to the acquisition of the same habit under normal conditions.

Their conclusion that the drug has a greater detrimental effect on a more difficult habit might explain the differences between our results and those of Siegel (**13**), and of Stainbrook and Löwenbach (**15**). Siegel found no differences between members of an electroshock group and controls on the very simple Graham-Gagné runway. Stainbrook and Löwenbach found that electricity produced no differences in the subsequent relearning of a simple water maze. It will be remembered

that the Lashley III maze used in our experiment was considered a difficult instrument by Lashley (**5**).

Riess' and Berman's theory that cerebral anoxia is effective in destroying more recent habits can also be applied to this experiment. Wortis, Shaskan, Impastato, and Almanski (**17**) found markedly impaired respiration of brain tissue following single and repeated electric shocks in rats. Cerebral anoxia, then, might be the mechanism operating to destroy the maze habit in our experiment.

The present experiment permits us to make no statement about the duration of the deficit following electroshock. Our rats started relearning the day after the last of the shocks. A parallel study by McGinnies and Schlosberg (to appear later), utilizing double alternation lever pressing, suggested that rats recover from the effects of single electroshock after about five days. It will be remembered that some of the clinical studies suggest that the memory loss is transitory. Eventually we hope to run a group of rats who start relearning a week after the last electroshock. But our present results show that there is a clear-cut deficit immediately after a series of 30 electroshock convulsions.

V. Summary and Conclusions

1. Twenty-eight male white rats were trained on the Lashley III maze until they had reached a criterion of three errorless trials.

2. Twelve of these animals, making up the Experimental Shock Group, then received one daily cerebral electroshock for thirty days. The shock consisted of 85 volts A. C., lasting .22 sec. Eight animals, the Shock Control Group, received a similar series of shocks through the back legs. The remaining eight rats, the Retention Control Group, were simply left in their cages during this period.

3. Immediately after the end of the 30-day period all rats were retrained on the maze until they had returned to the original criterion. Two of the Experimental Shock Group and one of the Shock Control Group died during the electroshock period, reducing the groups to ten and seven respectively.

4. The Experimental Shock Group was significantly inferior to the other two groups; (*a*) in terms of trials, errors, and time during relearning; (*b*) in terms of percentages saved in errors, trials, and time in relearning over learning; and (*c*) in terms of errors and time on the first retention trial. The Shock Control Group was slightly, but not significantly, inferior to the Retention Control Group. These results suggest that shocks through the head cause an impairment of the central nervous system that shows up as either a loss of retention or of relearning ability (or of both) for this difficult habit. Most of the loss seems to be in retention.

5. We believe that the loss of the habit and probably also the curing of mental patients are due, not to the punishment or fear of electroshock, but to a direct effect on cerebral tissue. A possible hypothesis as to the physiological action of electricity is presented.

(Manuscript received December 8, 1944)

Bibliography

1. Bunch, M. E., & Mueller, C. G. The influence of metrazol upon maze learning ability. *J. comp. Psychol.*, 1941, **32**, 569–574.
2. Graham, C. H., & Gagné, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
3. Heron, W. T., & Carlson, W. S. The effects of metrazol shock on retention of the maze habit. *J. comp. Psychol.*, 1941, **32**, 307–309.
4. Kessler, M., & Gellhorn, E. The effect of electrically and chemically induced convulsions on conditioned reflexes. *Amer. J. Psychiat.*, 1943, **99**, 687–691.
5. Lashley, K. S. *Brain mechanisms and intelligence.* Chicago: Univ. Chicago Press, 1929. Pp. 186.
6. Levy, N. A., Serota, H. M., & Grinker, R. R. Disturbances in brain function following convulsive shock therapy. *Arch. Neurol. Psychiat.*, 1942, **47**, 1009–1029.
7. Loken, R. D. Metrazol and maze behavior. *J. comp. Psychol.*, 1941, **32**, 11–16.
8. Page, J. D. Studies in electrically induced convulsions in animals. *J. comp. Psychol.*, 1941, **31**, 181–194.
9. Rennie, T. A. C. Present status of shock therapy. *Psychiatry*, 1943, **6**, 127–137.
10. Riess, B. F., & Berman, L. The mechanism of the insulin effect on abnormal behavior. *Amer. J. Psychiat.*, 1944, **100**, 674–680.
11. Rosen, V. H., & Gantt, W. H. Effect of metrazol convulsions on conditioned reflexes in dogs. *Arch. Neurol. Psychiat.*, 1943, **50**, 8–17.
12. Sherman, I., Mergener, J., & Levitin, D. The effect of convulsive treatment on memory. *Amer. J. Psychiat.*, 1941, **98**, 401–403.
13. Siegel, P. S. The effect of electroshock convulsions on the acquisition of a simple running response in the rat. *J. comp. Psychol.*, 1943, **36**, 61–65.
14. Smith, L. H., Hughes, J., Hastings, D. W., & Alpers, B. J. Electroshock treatment in the psychoses. *Amer. J. Psychiat.*, 1942, **98**, 558–561.
15. Stainbrook, E. J., & Löwenbach, H. The reorientation and maze behavior of the rat after noise-fright and electroshock convulsions. *J. comp. Psychol.*, 1942, **34**, 293–299.
16. Wortis, J. Physiological treatment of the psychoses. In: Review of Psychiatric Progress 1942. *Amer. J. Psychiat.*, 1943, **99**, 602–604.
17. Wortis, S. B., Shaskan, D., Impastato, D., & Almanski, R. Brain metabolism, VII: The effects of electric shock and some newer drugs. *Amer. J. Psychiat.*, 1941, **98**, 354–359.
18. Zubin, J., & Barrera, S. E. Effect of electric convulsive therapy on memory. *Proc. Soc. exp. Biol. Med.*, 1941, **48**, 596–597.

HEART RATE AND FREQUENCY OF BLINKING AS INDICES OF VISUAL EFFICIENCY[1]

BY M. E. BITTERMAN

Cornell University

Introduction

In a recent exchange of opinion, Luckiesh (**16**) and Tinker (**38, 39**) have argued the merits of heart rate, frequency of blinking, and speed of performance as indices of the ease of reading. This discussion brings into focus the chief disagreements of leading investigators in the field of visual efficiency.

Paterson and Tinker have emphasized the desirability of increasing the speed of reading and have adopted this criterion for the evaluation of typographical practices (**28**) and illumination requirements (**35, 36**). Their specific recommendations are based upon studies with a standardized modification of the Chapman-Cook Speed of Reading Test (**7**) which provides alternate forms of comparable difficulty and a simple cross-out technique for checking on comprehension.[2] Paterson and Tinker have, however, neglected the problem of minimizing discomfort and fatigue in reading, assuming that these factors do not vary independently of the performance index. The following paragraph from *How to make type readable* indicates their position with respect to this problem:

> "Fortunately, speed factors can be introduced into printing without sacrificing ease and comfort. Readibility of print as influenced by size of type, style of type face, quality of printing surface and other factors will not only prevent or reduce fatigue and feelings of eye strain but may also facilitate the speed with which the printed material can be assimilated. Hence speed tests can generally be relied upon to check the efficiency of traditional typographical practices as well as innovations" (p. 2).

On the basis of extensive research, Luckiesh and Moss (**20, 21**) have directly contradicted these conclusions. Despite the fact that Paterson and Tinker have reported significant variations in speed of reading consequent to fairly limited typographical changes, Luckiesh

[1] This paper is an adaptation of Part I of *Studies in Visual Fatigue and Efficiency*, a doctoral dissertation presented to the Graduate School of Cornell University. The writer is indebted to Professor T. A. Ryan under whose direction the work was performed. Thanks are due the Carrier Corporation of Syracuse, N. Y., for its generous gift of the air-conditioning apparatus required for the experiments, and the Psychological Corporation of New York City for permitting the reproduction of the Minnesota Vocational Test for Clerical Workers.

[2] Many of the speed studies were supplemented with photographic analysis of eye-movements (**27, 29, 40**).

and Moss have taken the position that the performance index is insufficiently sensitive. They believe that habits of reading and central mechanisms of compensation operate to maintain relatively constant output except under the least adequate conditions of work or at the extremes of fatigue. When level of performance is maintained in spite of unfavorable conditions, however, the 'cost' of the work to the organism is increased, and it is this expenditure of 'resources' which Luckiesh and Moss have attempted to measure. Their search for suitable indices revealed a number of involuntary, 'psychophysiological' processes—heart rate, frequency of blinking, and peripheral muscular tension—which vary continuously with the duration of reading. Assuming that fatigue increases with working time, and that the expenditure of effort required to maintain adequate performance increases with fatigue, Luckiesh and Moss concluded that these indices would serve as measures of readability. Additional evidence is provided by the observation that the same functions are affected by the adequacy of conditions of work as estimated from independent criteria. Glare, for example, which is known to be a deleterious factor in seeing, augments rate of blinking.[3]

Speed of reading.—The difficulties of performance indices are well known (**33**). McNally (**24**) has recently reviewed the literature and reported that "in general, investigators seem to agree that work output as a criterion of readibility is not so sensitive as could be desired" (p. 14). The significant findings of Paterson and Tinker (**28**) cannot, however, be ignored, and it is likely, as Tinker has pointed out (**38**), that the failure of Luckiesh and Moss to obtain similar results may be traced to the use of unstandardized material, failure to check on comprehension, and instruction of subjects to read at 'natural rate.'[4] On the other hand, studies of efficiency must, by definition, take account of the cost of work, and for this reason the promising findings of Luckiesh and Moss should be given careful consideration.

Indices of cost.—Of the three major indices of cost proposed by Luckiesh and Moss, peripheral muscular tension is most readily acceptable (**17, 18**). A number of other investigators (**8, 9, 11, 12**) have reported that tension increases with the duration of mental work and

[3] In one experiment (**21**) the production of artificial refractive errors by the use of lenses increased blink rate in 28 of 30 *Ss*. When one of the exceptional cases was examined by a refractionist it was found that an additional correction of the kind used in the experiment was required by the *S*.

[4] Actually Luckiesh and Moss have on occasion studied maximum rate (**21**). At first thought it may appear that the study of reading at the 'natural rate' provides a more realistic investigation. Pyke (**32**) considered the problem in his pioneer investigation and decided that it was necessary to sacrifice some degree of realism for the purpose of more adequate experimental control. There is, however, no single natural rate. Reading behavior varies widely, both qualitatively and quantitatively, with the material at hand and with the purpose of the reader (Vernon, **41**).

that the level of tension is a function of the severity of the task.[5] Freeman (**10**) has proposed the following explanation of related findings:

"Under such categories as distraction, fatigue and effort are usually grouped the conditions which disturb the precision of behavior and the means by which the organism meets these difficulties. The tendency to compensate for or overcome resistances blocking specifically integrated acts has been noted by many authors. It is as if the organism has standards which it strives to maintain. Such maintenance requires greater or less energy expenditure dependent upon circumstance. Increases in muscular tension apparently act to reinforce reaction systems already in progress in their competition with incompatible tendencies" (pp. 443–444).

With respect to the other indices—heart rate and frequency of blinking—there have, however, been a number of contradictory findings. It is with the validity of these functions that the present investigation is primarily concerned.

According to Litinski (**15**), pioneer work with the blinking index was done by Fick and later by Katz toward the end of the 19th century.[6] Fick reported that the frequency of blinking varies inversely with type size and is augmented by glare. Katz found that rate of blinking was increased in two *S*s when artificial refractive errors were induced through the use of prisms. Self-observation led him to conclude that blink rate is a function of the difficulty of the conditions of visual work. Litinski himself studied the relation of blinking to 'density' of type [7] but did not get consistent results. He concluded, however, that the index is adequate as a measure of 'general' fatigue (irrespective of the ocular burden), and he suggested its use in studies of the length of the working day.

On the basis of the conclusion of Ponder and Kennedy (**31**) that blinking may be related to 'mental tension' in that it may provide an outlet for 'nervous energy,' and on the basis of Litinski's work, Luckiesh and Moss (**20, 21**) began an extensive series of investigations of the frequency of blinking during reading. Their procedure—counting the lid movements as the *S* worked—was simple, and their results, to judge from their reports, were quite reliable. Blink rate was found to increase in the course of reading, the increments being inversely proportional to the level of illumination. A large number of other factors were similarly found to influence the response—type

[5] Recently Jacobson and Kraft (**13**) failed to obtain consistent results in the measurement of potentials from the *quadriceps femoris* during reading. The technique, which involved the direct insertion of needle-electrodes into the muscle, may have disturbed the inexperienced subjects and obscured the effects of work. That no instructions concerning the expected level of performance (speed) were given may be a crucial factor. The findings reported below indicate, for example, that cardiac activity during reading varies with the nature of the instructions to the subject.

[6] Litinski does not make specific reference to the sources of this information.

[7] Density was measured by number of letters per unit area of the page. The nature of the reading material was not held constant.

face, size, and boldness; glare; contrast; brightness-ratio (brightness of the reading material in relation to that of the surroundings); color of paper; leading; line-length; and refractive errors of the eye.

A number of recent investigations have, however, cast some doubt on the validity of these findings. McNally (**24**) compared six variations of type size and form and could find no differences in readability on the basis either of blink rate or speed of reading. McFarland, Holway and Hurvich (**22**) reported that rate of blinking does not uniformly increase during reading, nor is it consistently related to level of illumination. They found a greater blink rate during the reading of a book than during the reading of microfilm even though the Ss preferred the former task.[8] These writers report a personal communication from R. L. Thorndike who measured blinking during five-min. periods before and after a day of close work in a printing establishment. Thorndike found that individual differences were marked and consistent, but that a decrease in frequency was as likely as an increase after the day's work. McFarland and his co-workers concluded that, since some Ss show decreases in rate during visual work and others show no change at all, blinking does not provide an adequate index of visual fatigue. MacPherson (**25**) recently has reached similar conclusions.

Luckiesh and Moss (**19, 20, 21**) have also emphasized the value of heart rate as a criterion of readability. They found that heart rate decreases in the course of reading, the degree of change being inversely proportional to the intensity of illumination. Many of the early workers (**2**) reported increases in pulse rate at the beginning of mental work and a decline in rate with continued activity, but, since Luckiesh and Moss did not present data on pre-work levels, it is impossible to say whether their findings represent a falling-off following an initial increment. In a review of the literature, Bills (**4**) has concluded that heart rate increases slightly during mental work when a resting level is achieved before the work is begun. It is conceivable, however, that the adjustment for near vision involved in reading invokes the oculo-cardiac reflex mechanism as a result of the increase in intra-ocular pressure which may stem from this adjustment (**34**).

[8] They did find, however, that voluntary 'relaxed' blinking during visual work forstalled certain changes which were considered valid indices of fatigue—dilation of the pupil and recession of the near-point. Bates (3) and Peppard (30), as well as Luckiesh and Moss (20, 21), have also considered the possible prophylactic function of blinking. Ponder and Kennedy (31) suggested that blinking may allow alternation in the tension of the ocular muscles to occur and thus eliminate early fatigue. Whether there is necessarily any contradiction involved in regarding blinking both as an index of fatigue and a preventative of fatigue is not immediately apparent. Certainly peripheral muscular tension provides an analogous case as the passage from Freeman quoted above indicates. Tension is regarded as an index of increased difficulty of work consequent to the development of fatigue and, at the same time, as a facilitative factor in performance.

Two studies of heart rate during visual work have failed to confirm the findings of Luckiesh and Moss. Atkins (**1**) found increments in pulse rate after one hour of work at various visual tasks. McFarland, Kner, and Berens (**23**) did report a general tendency for heart rate to decrease slightly during two-hour reading periods, but the decrement did not change with level of illumination (1 vs. 50 foot-candles). They point out that their experimental conditions provided for control of environmental temperature and suggest that this difference may account for the discrepancy between their results and those of Luckiesh and Moss. Luckiesh (**16**), on the other hand, has emphasized McFarland's use of subjects in the basal state, suggesting that basal heart rate may be less sensitive to conditions tending to produce decrement. A difference between the two experiments which neither set of investigators has mentioned lies in the fact that McFarland and co-workers used the manual method of pulse rate determination, while Luckiesh and Moss employed a cardiotachometer. Palpation may tend to increase the heart rate, at least momentarily, and thus obscure differential response to illumination.

Because of these contradictory findings and the importance of deriving valid criteria of the cost of visual work, it was considered worthwhile to conduct further studies of heart rate and frequency of blinking.

Experiment I: Effect of Type Size

The first experiment was designed to provide data on the variation in heart rate, frequency of blinking, and level of performance with

TABLE I

Plan of Experiment I

Day *						
1	A	B	L	B	S	B
2	A	B	S	B	L	B
Time (min.)	10–15	5	15	10	15	5

Key: A, adaptation period, occupied principally with attachment of electrodes and instruction of subject. Recording was begun directly following this period.
B, rest interval.
L, work on large type.
S, work on small type.

* For half of the Ss the daily order of procedure was reversed.

size of type. Short periods of work were employed since the concern was with level of difficulty rather than fatigue.[9]

[9] An index of effort may be employed to measure the latter aspect on the assumption that fatigue implies increased cost of work (**5**).

Procedure.—The Minnesota Vocational Test for Clerical Workers [10] was photographically reproduced in two sizes. Each edition was printed on non-glossy white paper of good quality and contained two complete test forms. The larger type (*L*) was approximately 10 points in size, while the smaller (*S*) was about 6 points. The *S*'s task was to check the paired names and number-groups for identity and non-identity and to indicate these judgments by marking a 'plus' or 'minus' with a pencil in the blanks provided. Responses were scored right-minus-wrong.

Each *S* served for 50 min. on each of two successive days. The order of work is given in Table I. After attachment of electrodes and instruction of the *S*, the recording equipment (see below) was switched on. The *S* rested for five min. and then worked for 15 min. on the tests in large type (*L*). Following this work there was a 10-min. rest period after which the *S* worked on small type (*S*) for 15 min. The session was concluded with a final rest period of five min. in

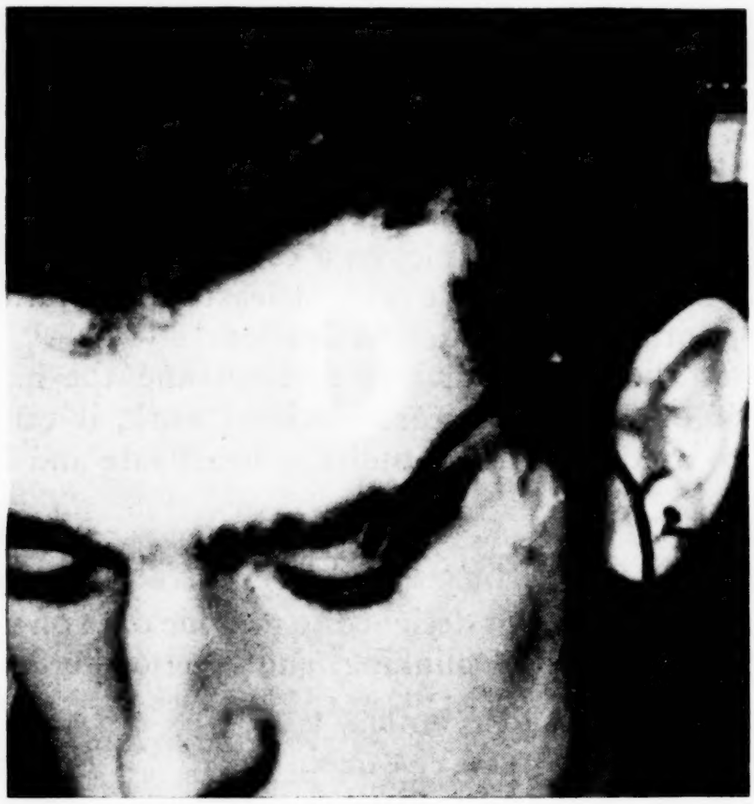


FIG. 1. Placement of electrodes for electromyographic recording of eyelid movements.

length. On the next day the procedure was repeated with the order of work reversed, i.e., the *S* worked first on (*S*). The alternation order controlled cumulative practice effects. By combining the results for the two days, level of performance on (*L*) could be compared with performance on (*S*). As a further control, the *daily* order of procedure was reversed for half of the *S*s, i.e., half of the *S*s began with (*L*) on the first day while the others began with (*S*).

The *S*s were specifically instructed to avoid being influenced by the obvious difference in type size. They were told that, since time limits for the two forms would differ and since a variety of factors would affect their performance, they had no basis for predicting the relative amounts of work which could be accomplished. They were asked to disregard all such considerations and to work as rapidly and as accurately as possible at all times.

Heart rate and frequency of blinking were measured continuously during the experimental sessions. Eyelid movements were recorded graphically by an improved technique (**6**, **14**). Action-potentials from the superior orbicular muscle of the non-dominant eye were led through push-pull resistance-capacity coupled amplifiers to a Grass ink-writer. The active electrode was

[10] Developed by Andrew and Paterson and published by the Psychological Corporation, New York City.

taped just below the brow in a manner which neither interfered with vision nor produced discomfort. A reference electrode was clipped to the lobe of one ear. The electrode placement may be seen in Fig. 1. Potentials arising from the heart (lead III) were amplified and recorded in a similar manner. The *S* sat in a comfortable chair which was provided with an adjustable writing board. Figure 2 illustrates the type of graphic record which resulted.

The level of illumination was three foot-candles at the writing-boards and remained constant throughout the experiment. At least 15 min. were allowed for adaptation to the illumination, the necessity for this procedure having been demonstrated by Tinker (37). The *S*s were specifically instructed not to close their eyes during the 'rest' periods. In order to approximate the conditions of Luckiesh and Moss, the *S*s were not in the basal state. They were, however, instructed to come to the laboratory at least two hours after a light meal.

A Carrier C–B Weathermaker afforded reliable control of temperature and humidity, factors which may affect frequency of blinking (26) as well as heart rate. Periodic checks during the

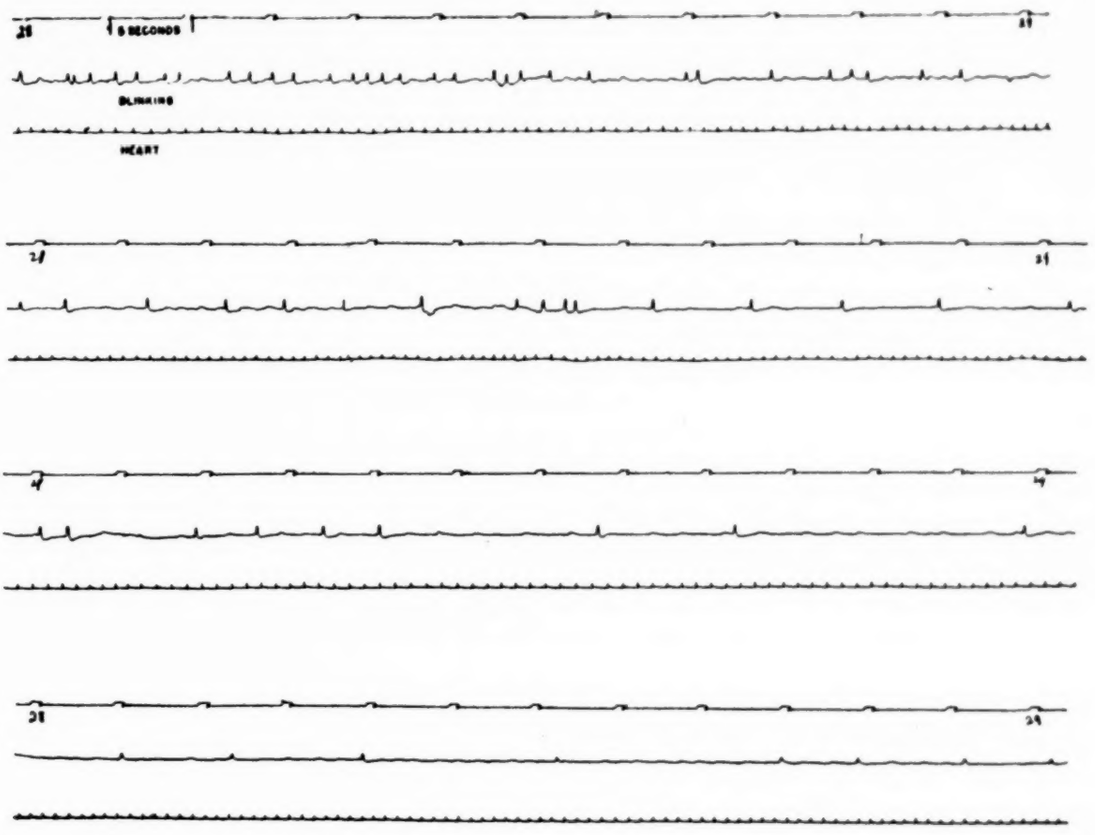


FIG. 2. Four records of blinking and heart rate during reading under low illumination in experiment II.

course of the experiments showed only minor fluctuations from the desired 67° Effective Temperature and 50 percent relative humidity. Dry-bulb readings ranged from 70° to 72° F. and wet-bulb readings from 59° to 61° F.

Fourteen undergraduate students in psychology served as *S*s. Eight were Naval trainees who satisfied the opthalmological standards set for officer-candidates. The others were women whose vision was either normal or corrected to normal, according to their own reports. They were informed only that various physiological processes would be measured while they worked.

Results.—For the 14 *S*s there was a significant relation between output and size of type. The mean score on (*L*) (representing 30 min. of work) was 523.6, while the corresponding mean for (*S*) was 487.4. The critical ratio of the difference was 5.1. There was, however, no reliable difference in accuracy of performance on the two forms. The

percentage of accuracy for (*L*) was 96.6 and for (*S*), 96.0. The critical ratio of this difference was only .8.

The heart rates recorded in this experiment support the conclusion of Bills (4). Table II shows that heart rate during the rest periods was significantly lower than during work on either form of the test. The mean for small type was higher than that for large type, although not significantly so (C.R. = 1.1).

Frequency of blinking [11] was not affected by size of type. The means for successive five-min. periods and for the total time of work are presented in Table III. These values are uniformly higher for

TABLE II

MEAN HEART RATES DURING REST AND WORK

Condition	Mean Heart Rate	Critical Ratios	
Rest	73.3	Rest vs. large type	6.7 *
Large type	79.3	Rest vs. small type	9.6 *
Small type	80.0	Large type vs. small type	1.1

* Significant at the one percent level.

TABLE III

MEAN FREQUENCY OF BLINKING IN RELATION TO SIZE OF TYPE

Type	Successive 5-Min. Periods			Total for 15 Min.
	1	2	3	
Large	26.8	32.5	28.6	87.9
Small	26.2	30.4	27.8	84.4
Difference	.6	2.1	.8	3.5
Critical Ratio	.5	1.3	.4	1.0

work with small type, although not significantly so. Variation in rate with working time may be seen more clearly from Fig. 3. It will be noted that the trends for the two conditions of work are quite similar. In each case a maximum occurs in the middle period. The values for this period are significantly higher than those for the first period but not significantly higher than the corresponding values for the third. In neither case is the frequency of blinking during the third period significantly greater than during the first.

Questioned at the conclusion of the experiments, all but one of the *S*s expressed a preference for the larger type. The exceptional *S* preferred small type on the basis of its greater 'compactness' but actually achieved a higher level of performance on (*L*).

EXPERIMENT II: LEVEL OF ILLUMINATION

The negative findings with respect to frequency of blinking obtained in the first experiment may be a function of the short periods

[11] Magnitude of blinking was ignored in the computations since there was very little intra-individual variation.

of work required or the set for speed given the *Ss*. The latter condition might also have been responsible for the increase in the heart rate during work. In a second study growing out of these considerations, the *Ss* read for longer periods under instructions which more closely resembled those in the experiments of Luckiesh and Moss. Level of illumination rather than type size was selected as a variable.

Procedure.—The *Ss* sat in a comfortable chair provided with a reading-board and read *The battle is the payoff* [12] for 40-min. periods on each of two successive days. The book was interesting but not exciting [13] and instructions were given to 'settle down and enjoy it.' Before and after the reading periods the *Ss* rested with eyes open for intervals of five min.

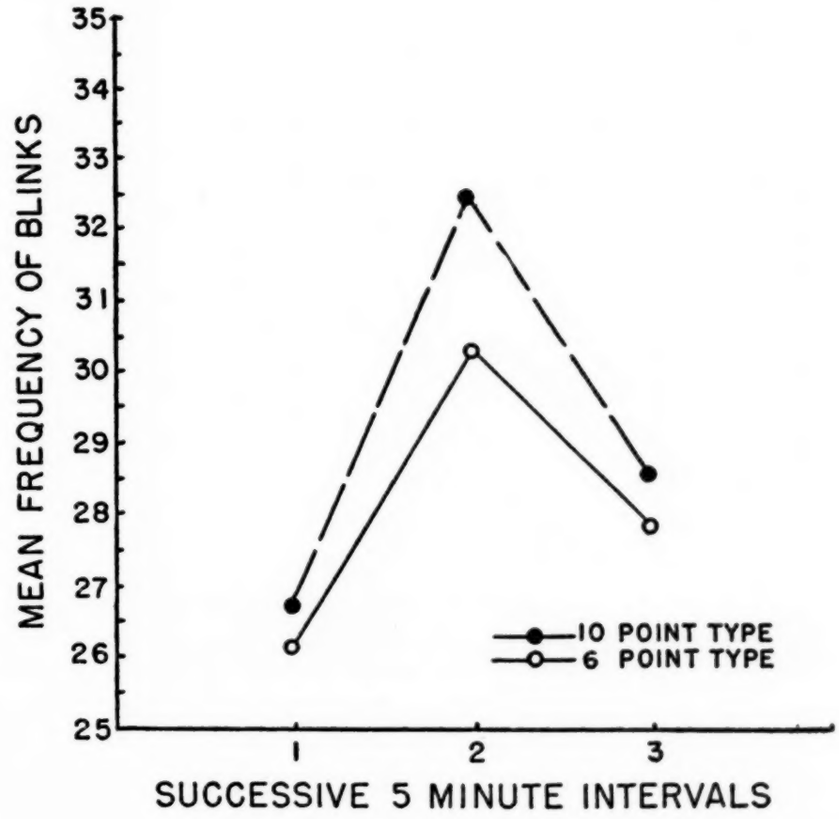


FIG. 3. Mean frequency of blinking during successive five-min. periods of work on each form in experiment I.

The book was printed in 10-point type on non-glossy paper. Half of the *Ss* read under illumination of three foot-candles on the first day and 91 foot-candles on the second. The others read first under the higher level. The illumination was diffuse and produced no appreciable glare. At least 15 min. were allowed for adaptation on each day (37). Temperature and humidity were maintained at the same levels as in the previous experiment. Continuous records of heart rate and frequency of blinking were made in the manner already described. Ten undergraduate students in psychology served as *Ss*. They were not in the basal state.

Results.—The reading periods were divided into eight five-min. intervals. The frequency of blinking during each interval and the

[12] By Ralph Ingersoll, Harcourt, Brace and Co., 1943.

[13] According to the judgment of the writer and the reports of the *Ss*. In no case was much more than one half of the book completed in the allotted time.

heart rate during the mid-minute of each interval were computed. Heart rate also was counted during the rest periods.

Exactly five of the 10 *Ss* preferred the low level of illumination; the remainder considered the higher level more comfortable.[14] For this reason two sets of comparisons were made—high versus low illumination and preferred versus non-preferred or 'alternate' illumination.

Mean frequencies of blinking during successive five-min. periods under various conditions are plotted in Fig. 4. None of the upward

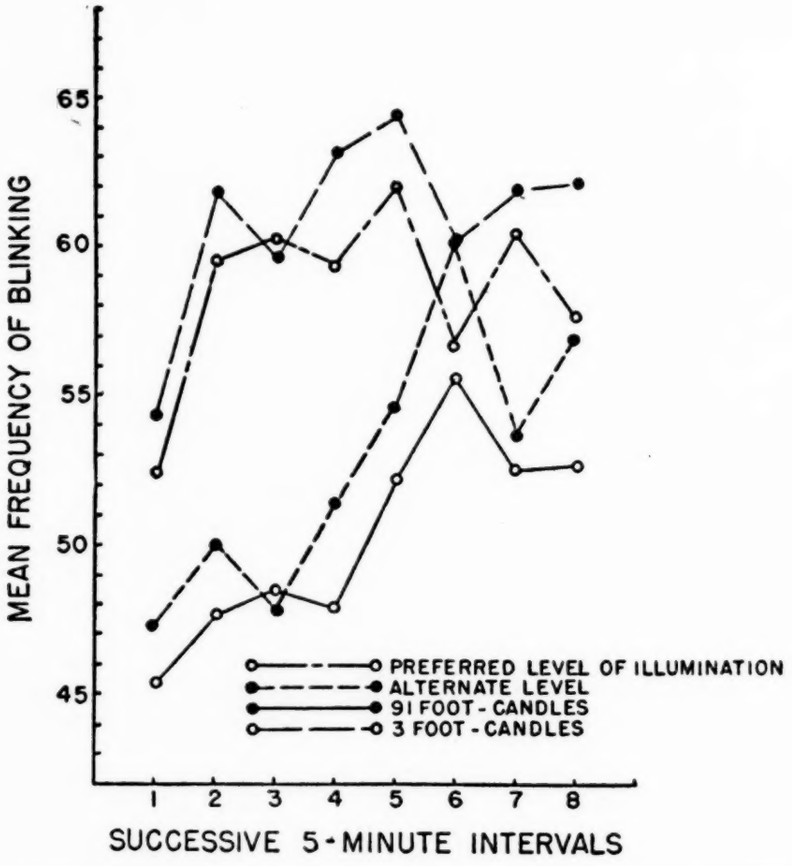


FIG. 4. Mean frequency of blinking during successive five-min. periods of reading under various conditions of illumination.

trends which the curves seem to show proved to be significant as a result of an analysis of variance. The *F*-values ranged from .63 to 1.92; $P > .05$. Individual curves presented in Fig. 5 show a great variety of trends. The total frequency of blinking under higher illumination actually was *greater* (C.R. = 2.7, $P < .05$) than under low illumination. Frequency of blinking under the preferred condition tended to be higher than under the alternate condition, although this difference was not significant (C.R. = 1.2).

Fig. 6 shows the changes in heart rate which occurred during

[14] The *Ss* were questioned at the conclusion of the final session. Preferences were not related to the level of illumination to which the *Ss* were adapted at the time.

reading under the different conditions. All of the curves show downward trends, but analysis of variance indicates that the trends for high illumination and non-preferred illumination are not significant ($F = 1.88, 1.41$; $P > .05$). Variance due to time of reading was, however, reliably greater than chance in the curve for low illumination ($F = 4.35$; $P < .01$) and in the curve for preferred illumination ($F = 2.66$; $P < .05$). Mean heart rates for high as compared with low illumination and for the preferred as compared with the alternate level tended to be higher, but in neither case was the difference significant (C.R. = 1.66, 0.4; $P > .05$). The curves did not show the

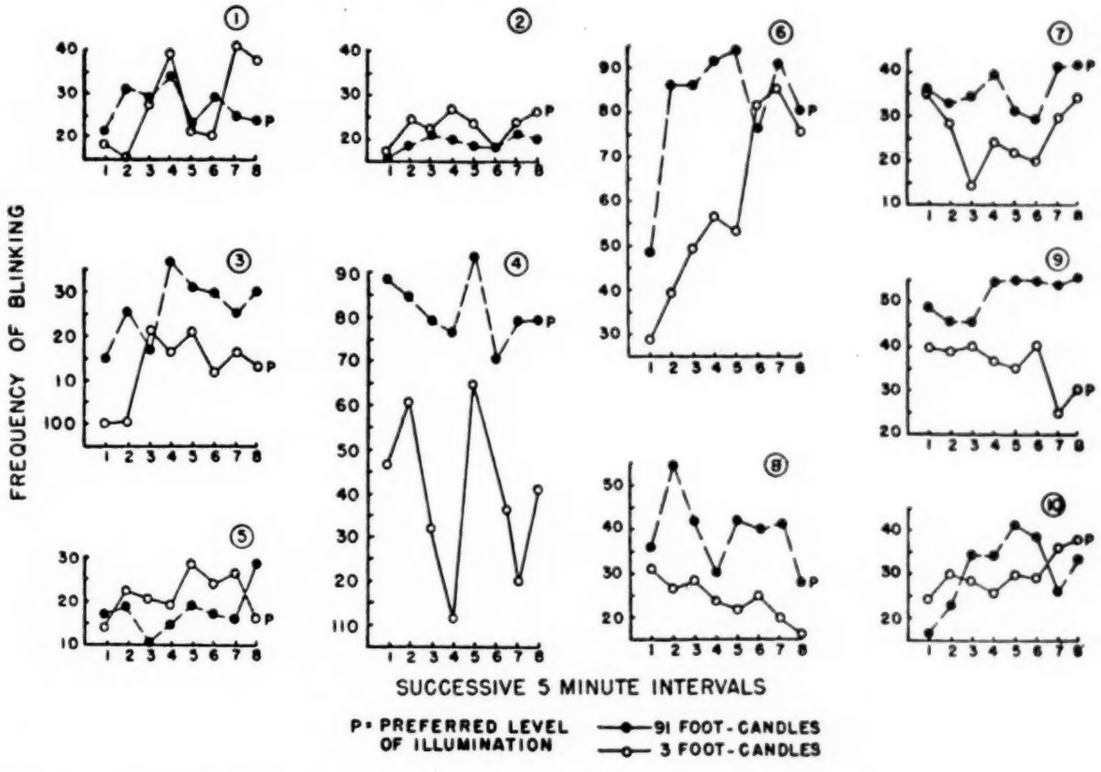


FIG. 5. Individual curves for the 10 *Ss*, showing frequency of blinking in successive five-min. periods of reading under the various conditions of illumination.

sharp rises at the termination of reading which were reported by Luckiesh and Moss (**21**).

DISCUSSION

It is clear that rate of blinking can have little value as an index of visual fatigue or effort. Although level of performance differed significantly for the two sizes of type employed in the first experiment, patterns of blinking were very similar under the two conditions. There was no tendency for blink rate to increase with the duration of work, but rather, for unknown reasons, a maximum frequency occurred during the middle periods of work. In the second experiment,

blink rate was actually higher during work under the more intense illumination, but was unrelated to *Ss*' preferences in the matter. Although frequency of blinking showed a tendency to increase with the duration of work under all conditions, none of these trends proved to be statistically reliable. The curves for individual *Ss* showed a great variety of changes with time, some rising, some falling, and some remaining level. Even if average trends had been significant, there-

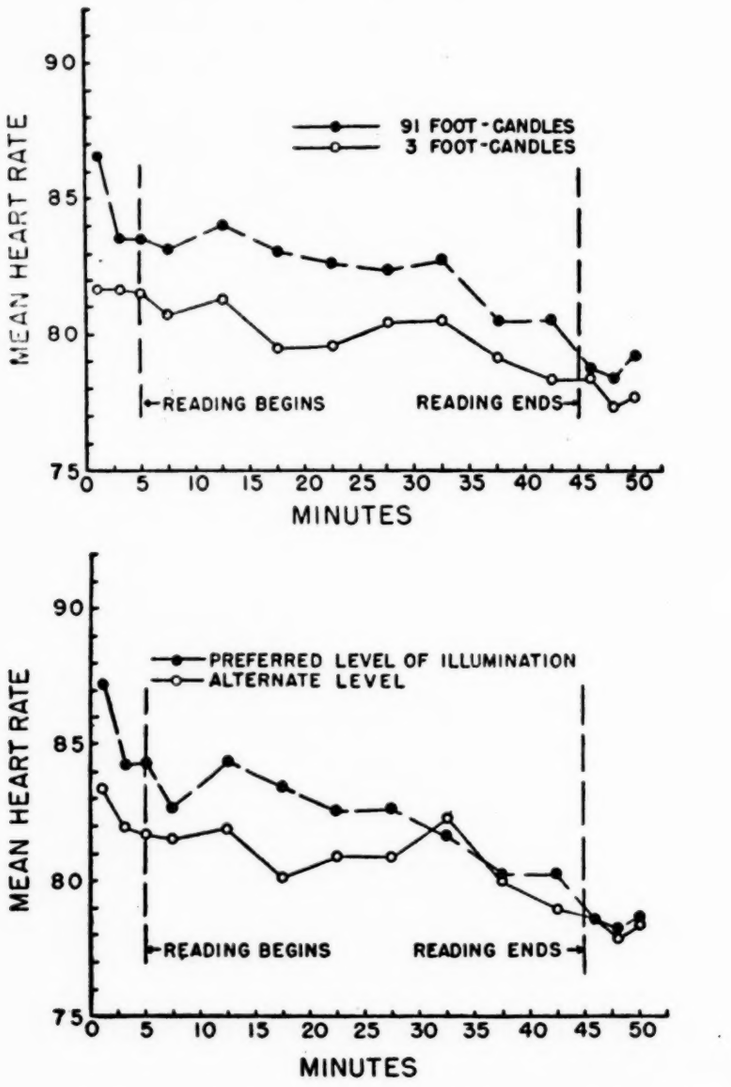


FIG. 6. Mean heart rate during rest and work under the various conditions of illumination.

fore, the individual variation would indicate that many factors play a part in determining the pattern.

With reference to heart rate, the present findings confirm those of Luckiesh and Moss in only one respect. During relaxed reading the heart rate shows a greater tendency to decrease with the passage of time under low as compared with high illumination. It is not justified, however, to assume that cardiac deceleration is a function of

increased expenditure of effort. It is well known that when a standing *S* assumes the sitting position a fall in heart rate occurs, and it is likely that the direct effect of intense visual stimulation would serve to counteract this tendency under the higher level of illumination. It should be noted that at the cessation of reading the heart rate did not rise as Luckiesh and Moss have claimed to demonstrate. Whatever the validity of this interpretation, the results of the first experiment make untenable the position that visual effort produces cardiac deceleration. Heart rate rose significantly when the *S*s were instructed to work as rapidly and as accurately as possible, and there was a tendency for the rate to be greater when the *S*s read the smaller type.

The findings of the first experiment support the position of Paterson and Tinker (**28**) that under suitably controlled conditions performance tests afford a reliable approach to problems of visual efficiency. Whether output indices are optimally sensitive cannot yet be decided, but at present we have little else on which to base practical recommendations regarding adequate conditions of work. It is recognized, of course, that studies of visual efficiency must, by definition, involve some measure of the cost of work. The results of other investigations indicate that further experiments with muscular tension might profitably be undertaken (**5**).

Summary

Continuous electromyographic recordings of cardiac and eyelid activity were made as *S*s performed visual tasks under a variety of conditions. The results do not support the conclusions of Luckiesh and Moss with respect to the value of heart rate and frequency of blinking as indices of the ease of visual work.

(Manuscript received January 19, 1945)

Bibliography

1. Atkins, E. W. The efficiency of the eye under different intensities of illumination. *J. comp. Psychol.*, 1927, **7**, 1–37.
2. Barmack, J. E. Boredom and other factors in the physiology of mental effort. *Arch. Psychol.*, No. 218, 1937.
3. Bates, W. H. *The cure of imperfect sight by treatment without glasses.* New York: Central Fixation Publishing Co., 1920.
4. Bills, A. G. Fatigue in mental work. *Physiol. Revs.*, 1937, **17**, 436–458.
5. Bitterman, M. E. Fatigue defined as reduced efficiency. *Amer. J. Psychol.*, 1944, **57**, 569–573.
6. Bitterman, M. E. Electromyographic recording of eyelid movements. *Amer. J. Psychol.*, 1945, **58**, 112–113.
7. Chapman, J. C., & Cook, S. The principle of the single variable in a speed of reading cross-out test. *J. educ. Res.*, 1923, **8**, 389–396.
8. Davis, R. C. The relation of action potentials to difficulty and frustration. *J. exp. Psychol.*, 1938, **23**, 141–158.

9. Davis, R. C. Muscular activity during 'mental work.' *J. exp. Psychol.*, 1939, **24**, 451–465.
10. Freeman, G. L. Mental activity and the muscular processes. *Psychol. Rev.*, 1931, **38**, 428–429.
11. Freeman, G. L., & Lindley, S. B. Two neuro-muscular indices of fatigue. *J. exp. Psychol.*, 1931, **14**, 567–605.
12. Hadley, J. M. Some relationships between electrical signs of central and peripheral activity: II. During 'Mental work.' *J. exp. Psychol.*, 1941, **28**, 53–62.
13. Jacobson, E., & Kraft, F. L. Contraction potentials (right quadriceps femoris) in man during reading. *Amer. J. Physiol.*, 1942, **137**, 1–5.
14. King, H. E., & Landis, C. A comparison of eyelid responses conditioned with reflex and voluntary reinforcement in normal individuals and in psychiatric patients. *J. exp. Psychol.*, 1943, **33**, 210–220.
15. Litinski, G. A. (Recording of blinking as a method of studying visual fatigue in children resulting from reading.) *Sovet. Vestnik Oftal.*, 1934, **4**, 275–286.
16. Luckiesh, M. Some comments on Dr. Tinker's review of 'Reading as a visual task.' *J. appl. Psychol.*, 1943, **27**, 360–362.
17. Luckiesh, M., & Moss, F. K. Muscular tension resulting from glare. *J. gen. Psychol.*, 1933, **8**, 455–460.
18. Luckiesh, M., & Moss, F. K. A correlation between illumination intensity and nervous muscular tension resulting from visual effort. *J. exp. Psychol.*, 1933, **16**, 540–555.
19. Luckiesh, M., & Moss, F. K. The effect of visual effort on the heart rate. *J. gen. Psychol.*, 1935, **13**, 131–139.
20. Luckiesh, M., & Moss, F. K. *The science of seeing.* New York: Van Nostrand, 1937.
21. Luckiesh, M., & Moss, F. K. *Reading as a visual task.* New York: Van Nostrand, 1942.
22. McFarland, R. A., Holway, A. H., & Hurvich, L. M. *Studies of visual fatigue.* Boston: Graduate School of Business Admin., Harvard Univ., 1942.
23. McFarland, R. A., Kner, C. A., & Berens, C. Metabolism and pulse rate as related to reading under high and low levels of illumination. *J. exper. Psychol.*, 1939, **25**, 65–75.
24. McNally, H. J. The readability of certain type sizes and forms in sight-saving classes. *Teach. Coll. Contrib. Ed.*, 1943, No. 883.
25. MacPherson, S. J. The effectiveness of lighting—its numerical assessment by methods based on blinking rate. *Trans. Illum. Engng. Soc.*, 1943, **38**, 520–522.
26. Newhall, S. M. The control of eyelid movement in visual experiments. *Amer. J. Psychol.*, 1932, **44**, 555–570.
27. Paterson, D. G., & Tinker, M. A. Influence of line width on eye movements. *J. exp. Psychol.*, 1940, **27**, 572–577.
28. Paterson, D. G., & Tinker, M. A. *How to make type readable.* New York: Harper, 1940.
29. Paterson, D. G., & Tinker, M. A. Eye-movements in reading optimal and non-optimal typography. *J. exp. Psychol.*, 1944, **34**, 80–83.
30. Peppard, H. M. *Sight without glasses.* Garden City & New York: Blue Ribbon Books, 1940.
31. Ponder, E., & Kennedy, W. P. On the act of blinking. *Quart. J. exp. Physiol.*, 1927–8, **18**, 89–110.
32. Pyke, R. L. *The legibility of print.* London: His Majesty's Stationery Office, 1926.
33. Ryan, T. A. Varieties of fatigue. *Amer. J. Psychol.*, 1944, **57**, 565–569.
34. Southall, J. P. C. *Introduction to physiological optics.* London: Oxford Univ. Press, 1937.
35. Tinker, M. A. The effect of illumination intensities upon speed of perception and upon fatigue in reading. *J. educ. Psychol.*, 1939, **30**, 561–571.
36. Tinker, M. A. Illumination standards for effective and comfortable vision. *J. consult. Psychol.*, 1939, **3**, 11–20.
37. Tinker, M. A. The effect of adaptation upon visual efficiency in illumination studies. *Amer. J. Optom.*, 1942, **19**, 143–151.
38. Tinker, M. A. Review of Luckiesh, M., and Moss, F. K., *Reading as a visual task.* *J. appl. Psychol.*, 1943, **27**, 116–119.
39. Tinker, M. A. A reply to Dr. Luckiesh. *J. appl. Psychol.*, 1943, **27**, 469–472.
40. Tinker, M. A., & Paterson, D. G. Influence of type form on eye movements. *J. exp. Psychol.*, 1939, **25**, 528–531.
41. Vernon, M. D. *The experimental study of reading.* Cambridge, Eng.: Cambridge Univ. Press, 1931.

HABIT STRENGTH AS A FUNCTION OF THE PATTERN OF REINFORCEMENT

BY O. H. MOWRER AND HELEN JONES

Harvard University

It was long assumed that the basic truth concerning the psychology of learning was embodied in the Law of Exercise (Repetition, Use, Frequency), which held that every performance of a habit necessarily strengthened that habit. But evidence gradually accumulated which suggested that 'exercise,' as such, has little, if anything, to do with the strength of a habit. It was rather the presence or absence of *reward* (Law of Effect) that seemed crucial, since only rewarded responses grow stronger with repetition and non-rewarded (or punished) responses are actually inhibited thereby. Even this proposition, that every rewarded repetition of a habit strengthens it, has, however, recently been questioned. It is our purpose to consider whether the doubt which has thus been thrown on this proposition is justified.

If it be assumed that reward is the essential factor in strengthening a habit and that non-rewarded repetitions have a weakening effect, it would seem to follow that habits which are rewarded everytime they occur would be strengthened more than habits which are rewarded only intermittently. But recent experimentation has brought the validity of this inference under suspicion. When 'resistance to extinction,' i.e., the extent to which a previously rewarded habit will continue to occur after reward is eliminated, is used as a measure of habit strength, it is found that a habit which has been developed on the basis of intermittent reward is, by this criterion, 'stronger' than a habit which has been continuously rewarded. The experiment here reported represents an attempt to explain this paradox and to show that the experimental facts are actually consistent with the Law of Effect.

Formulation of the Problem

Since the present investigation was prompted largely by Humphreys' research on intermittent vs. constant reinforcement, it is necessary to give a brief description of this work.

Humphreys performed three experiments with human subjects, in the first of which he used the conditioned eyelid response (**12**), in the second a verbal response (**13**), and in the third the conditioned psycho-

galvanic response (14). In all three experiments a conditioned stimulus was followed by the unconditioned stimulus 100 percent of the time in one group, whereas in a second group, the conditioned stimulus was followed by the unconditioned stimulus only 50 percent of the time, in a predetermined but unpredictable order (though never with more than two reinforced or two non-reinforced trials in succession).[1] In each experiment Humphreys found that the 50-percent-reinforcement group showed reliably greater resistance to extinction than did the 100-percent-reinforcement group.

Discarding Hull's 'excitatory potential' concept (10), Skinner's 'reflex reserve' theory (28), and Hovland's hypothesis of 'extinction of reinforcement' (8) as inadequate to account for his findings, Humphreys has put forward the alternative suggestion that "conditioned responses are a consequence of anticipated reinforcement, extinction [a consequence] of anticipated non-reinforcement and that the role of frequency in the repetition of reinforcement and non-reinforcement is by way of its influence on the subject's expectation of the stimuli which are to appear" (12, p. 150). Elsewhere, Humphreys says that the fact that discontinuous reinforcement produces greater resistance to extinction than does continuous reinforcement is "apparently due to difficulty in forming an hypothesis of continuous non-reinforcement after discontinuous reinforcement in acquisition" (14, p. 74).

More recent investigations by Humphreys (15) and Finger (5, 6), with rats as *S*s, have likewise shown greater resistance to extinction after discontinuous than after continuous reinforcement. But Brown [2] has questioned whether Humphreys' 'expectation' hypothesis is scientifically meaningful at the infra-human level and has proposed an alternative interpretation. Brown suggests that an understanding of the results obtained with intermittent reinforcement hinges on the definition of the word, 'response.' Let us suppose that a hungry animal has been trained with intermittent reinforcement to press a bar as a means of obtaining food, i.e., has sometimes been rewarded after pressing the bar once, sometimes after pressing it twice, sometimes after pressing it three times. The 'response' which is thus effective in securing food consists, therefore, sometimes of one, sometimes of two, sometimes of three individual bar depressions. During

[1] For Pavlov and many of his followers, the establishment of conditioned responses was assumed to involve nothing more than the paired presentation of the conditioned and the unconditioned stimulus. But it is now clear that even in avoidance conditioning of the kind employed in Humphreys' three experiments, the learning that occurs is subject to the same principles as control trial-and-error learning, where drive and reward play such an unmistakable role (18, 19, 22). Therefore, whenever we use the term 'reinforcement' we assume it to imply reward (drive reduction), even though it be in describing experiments in which the experimenter thought of reinforcement in purely Pavlovian (associationistic) terms.

[2] Dr. Judson S. Brown—personal communication.

extinction, animals trained to depress the bar more than once as a means of getting food might, consequently, be expected to depress the bar more times *but not necessarily to make more 'responses' as thus defined* than animals which, having been rewarded consistently rather than intermittently, have learned exclusively a 'response' consisting of *only one* depression of the bar. In other words, Brown's suggestion is that we may have been deceiving ourselves in this type of situation by equating the term 'response' to single bar depressions instead of interpreting it to mean the total sequence, or *pattern*, of behavior leading up to the reward.[3]

The foregoing considerations raise the following specific question: If one uses as the unit for measuring resistance to extinction the same pattern of behavior which served as the instrumental 'response,' or 'act,' during acquisition, will intermittent reinforcement still produce greater habit strength than does reinforcement that is continuous?

In order to answer this question it seemed desirable to carry out an experiment involving one group of rats in which reward would follow each bar depression, three groups whose response pattern, or unit, would always consist, respectively, of two, three, or four depressions of the bar, and a final group which would be rewarded not only discontinuously but also randomly, on the average after every two and one-half bar depressions.

The reason for this particular combination of groups is that prior experiments on intermittency of reinforcement have been to some extent confounded: groups of *S*s which have been rewarded continuously and, necessarily, regularly have been compared with groups which have been rewarded discontinuously *and irregularly*. This confounding leaves open the question whether the obtained results are due to one, the other, or both of these factors. In our experiment, Group I is rewarded regularly and continuously, Groups II, III, and IV are rewarded regularly but discontinuously, and Group V is rewarded discontinuously and irregularly. In this way we have a combination of groups which makes possible all the comparisons necessary to determine whether it is to discontinuity, irregularity, or both that the phenomenon in question is attributable.[4]

Apparatus

The apparatus used in this experiment has been fully described elsewhere (21). It consists, essentially, of a wooden box, approximately one foot square, with a glass front and a wire-mesh floor. Originally there were two ¼-in. brass rods which extended horizontally into the box on

[3] *Cf.* Murray's (24) similar emphasis upon the distinction between what he terms 'actones' and 'acts.'

[4] We are indebted to Mr. John L. Wallin for useful suggestions concerning the logic and design of this experiment.

either side of the food trough, which was located in the center of the interior rear wall; but for the purposes of this experiment, only the bar on the right was used.

In order to avoid differential auditory cues following bar depressions which were and were not to produce food, the automatic delivery mechanism employed in earlier experiments was disconected and the food (a small, cylindrical pellet of Purina dog chow) was manually released by *E* from the rear of the apparatus into the shute-like trough. However, the criterion as to when the bar had been pushed down sufficiently far to count as a full depression was standardized by a system of electrical contacts and a relay which caused a small light to flash on in *E*'s (but not in the rat's) field of vision whenever the bar was depressed by a predetermined amount. This electrical set-up was also arranged so that each flash of the light signified a full excursion of the bar, i.e., the bar could not be depressed until the light came on and then 'jiggled' as a means of making multiple flashes of light. Once the light had come on it would stay on until the bar was released and allowed to return to its normal position; only by means of another complete depression of the bar could the light be made to come on again.

A cumulative work recorder, which has also been separately described (**20**), was connected to the apparatus so that each time the light flashed a small increment of change was registered. Records thus obtained will be reproduced on a later page.

The reasons for arranging the apparatus so that only complete, discrete bar depressions would cause the light to flash and the recorder to operate are obvious, but the precautions taken to eliminate differential cues require a word of explanation. It was felt that if bar depressions which produced food resulted in a sound (due to the automatic delivery mechanism) which was noticeably different from the sound accompanying depressions which did not produce food, the animals would simply learn to go to the bar and pump away until they got the auditory cue signifying the reward. Similarly, in extinction there would have been a tendency to keep pushing the bar until the tell-tale sound occurred. It therefore seemed desirable to circumvent the entire problem by simply eliminating all differential cues. By releasing the pellet manually, it was possible to make the accompanying sound almost imperceptible; and by withholding the pellet until the animal came and looked in the trough, we prevented even this slight noise from serving as a cue as to whether any given bar depression would or would not be followed by food. The only way an animal could 'know' when he had pushed the bar enough, without actually going and looking in the trough, was on the basis of his own kinaesthesis.[5]

Procedure

Habituation and acquisition.—Thirty Lashley-strain male rats (four months old) were reduced and held for the duration of this investigation to 15 percent below normal body weight. At the outset of the experiment, each animal was given 20 pellets of food in the apparatus for purposes of habituation. On the first training day, the bar (requiring a minimum pull of five grams to depress it) was inserted into the apparatus, and the animals were required to push it in order to obtain the food. Each animal was given 20 trials. This same procedure was repeated on the second and third days. On the fourth day the animals made the first five responses with the bar requiring a downward pull of five grams as usual, but for the next five responses the bar was weighted so as to require a pull of 15 gm., and on the final 10 responses the bar required a pull of 30 gm. During the remainder of the experiment the bar was kept weighted so as to require this latter amount of downward pressure to operate it.[6]

---

[5] Earlier unpublished research indicates that rats can learn to discriminate between one and two bar depressions and between two and three, but beyond this their discrimination becomes unreliable. In the present experiment, multiple bar depressions often occurred between inspections of the food trough, but the more usual procedure in all groups was to press, look; press, look; etc.

[6] The bar was made easy to depress during the early phase of training and reward was provided after each response in order to get the bar-pressing habit established more quickly and more reliably than would have been possible if the various conditions employed in the later part of the experiment had existed from the beginning. Our reason for increasing the effort required to depress the bar on the fourth day of training was that we wanted this act to be the major element in the total sequence of behavior involved in going back and forth between the bar and the trough, eating the food, etc.

On the fifth day, after the preliminary training just described, the animals were randomly divided into five groups. The animals in Group I were rewarded after every bar depression, those in Group II after every second bar depression, those in Group III after every third, those in Group IV after every fourth, and those in Group V in random order, after one, two, three, or four bar depressions but, on the average, after two and one-half. These different procedures were followed from the fifth through the eleventh day, the animals in each group being given 20 reinforcements (pellets) daily, as on days one to four. By the end of the eleventh day each animal had thus received 220 reinforcements, the first 80 of which had been received under the same (preliminary) conditions for all groups, and the last 140 under conditions which were distinctive for each group.

Extinction.—Extinction began on the 12th day and continued through the 13th and 14th days. The animals were left in the apparatus for 20 min. on each of these days, free to press the bar at will but given no food whatever. (Outside the experimental situation, they received enough food to keep them at the standard body weight mentioned previously.) The number of bar depressions made during this period was counted by the observer and recorded automatically.

Results and Interpretations

Groups I, II, III, and IV will be discussed first, inasmuch as the problem investigated with the animals comprising these four groups is that of the effect of continuous reinforcement (Group I) vs. discontinuous, or intermittent, reinforcement (Groups II, III, and IV). Group V poses the separate question of the importance of irregularity, or randomness, as opposed to regularity of reinforcement (Groups I–IV).

Continuous vs. Discontinuous Reinforcement

Table I gives the average number of bar depressions made by Groups I, II, III, and IV during the three days of extinction. Fig. 1

TABLE I

Average Number of Bar Depressions Made on Three Successive Days of Extinction by Groups I, II, III, and IV

The bottom line (Totals) gives the average number of bar depressions made by each group during all three extinction days.

		Groups			
		I	II	III	IV
Extinction Days	I	77.4	105.5	113.8	155.1
	II	34.8	60.5	74.3	75.0
	III	15.8	22.0	27.4	42.2
	Totals.....	128.0	188.0	215.5	272.3

shows the actual extinction performances on Day I and Day III of typical animals in Groups I and IV. These records for individual animals give a graphic equivalent of the corresponding group averages appearing in the four corners of Table I. From Table I it is evident (*a*) that all groups made progressively fewer bar depressions on the three days of extinction and (*b*) that the animals in Groups I to IV

made progressively more bar depressions on each of the three days of extinction, as well as for the three extinction days considered as a whole (see Totals). From Fig. 1 it is further evident (*c*) that, as has repeatedly been noted by other investigators (**7**, **11**), the rate of bar-pressing activity was greatest at the beginning of each extinction session and was thereafter negatively accelerated.

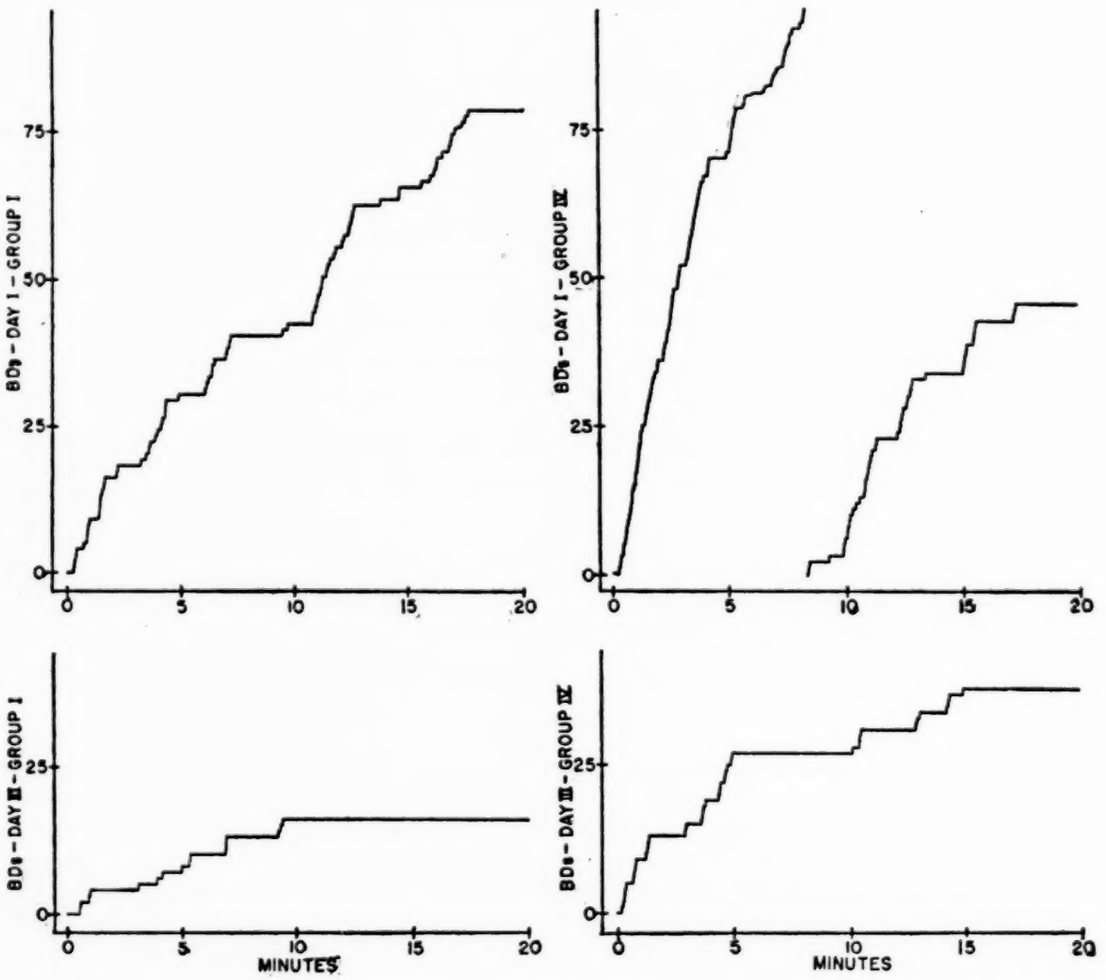


FIG. 1. Graphic recordings of the performance of a typical animal in Group I and a typical animal in Group IV on Day I and Day III of extinction. These four individual recordings, representing 79, 16, 143, and 38 bar depressions (BDs) respectively, correspond to the group averages appearing in the four corners of Table I. The two curves in the upper right-hand recording should be thought of as continuous (*cf.* **20**).

The three sets of curves appearing in Fig. 2 present these same data in a somewhat different manner. They represent the average number of bar-depressions made by each of the four groups, by two-min. periods, during each of the three 20-min. extinction sessions. The three trends mentioned in the preceding paragraph are again evident in Fig. 2.

The relationship between resistance to extinction and the reinforcement ratio used during acquisition becomes especially clear in

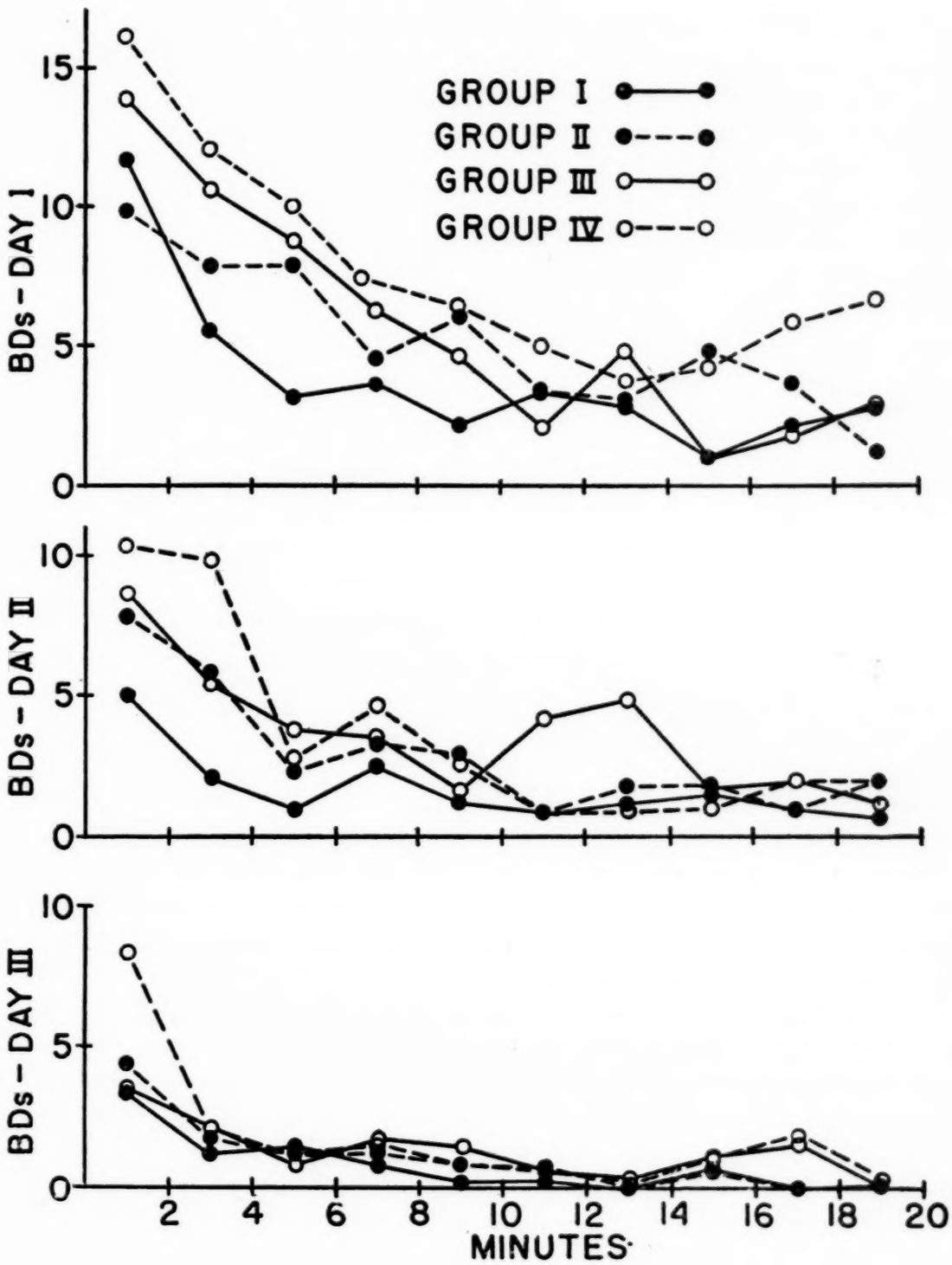


FIG. 2. Curves showing the average number of bar depressions (BDs) made by Groups I-IV on the three days of extinction. Each point on the curves represents the average number of bar depressions made in a two-min. interval.

Fig. 3. The broken line shows (*cf.* Totals in Table I) the average number of bar depressions made during all three extinction days by the animals in Groups I, II, III, and IV. This number was 128.0 for Group I, 188.0 for Group II, 215.5 for Group III, and 272.3 for

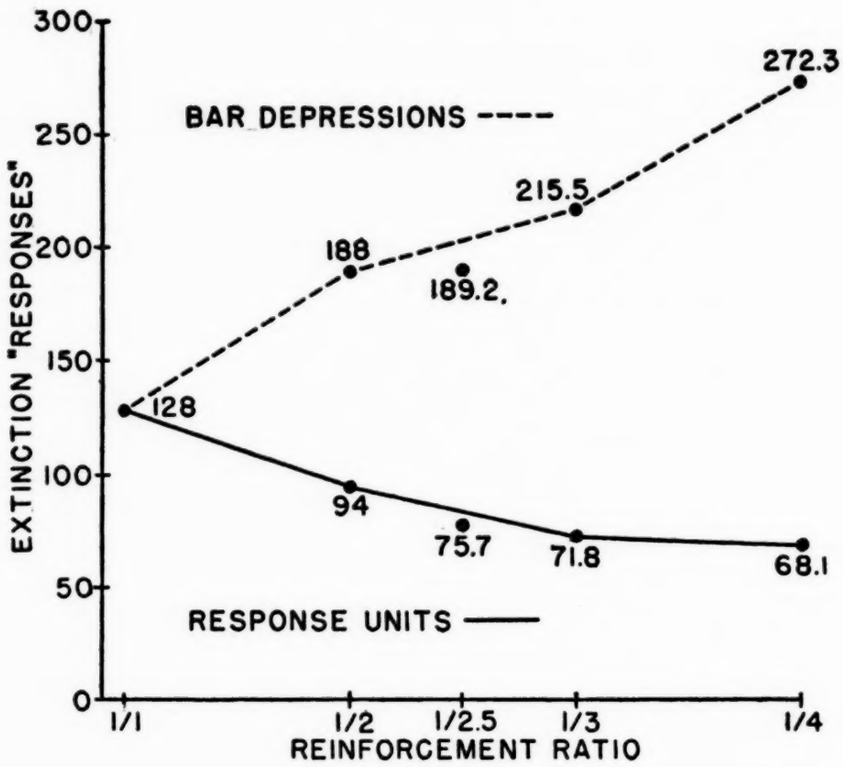


FIG. 3. Curves showing the average number of bar depressions as contrasted with response-units made by Groups I-IV (reinforcement ratio of 1/1 to 1/4) on all three days of extinction.

Group IV. An analysis of variance gives an F of 7.18 for the differences between these groups due to frequency of reinforcement, an F of only 4.43 being necessary for significance at the .01 level. The t-test shows significant differences at the .02 level or better between the mean of Group I and the means of all other groups and between the means of Groups II and IV.

From these results it would appear, in keeping with earlier findings, that not only is intermittent reinforcement reliably more effective than continuous reinforcement in establishing a tendency to depress the bar in the absence of reinforcement; it may also be said that, within the limits of this experiment, the functional relationship between these two variables is virtually linear.

Since the animals in Group I were rewarded during acquisition every time they pressed the bar, whereas the animals in Groups II, III, and IV received reinforcement only one-half, one-third, and one-fourth as often—and yet developed progressively greater apparent habit strength, as measured by resistance to extinction, it is easy to see how results of this kind might indeed seem incompatible with the assumption that every rewarded repetition of a habit strengthens it and non-rewarded repetitions weaken it. There are, however, at least two ways in which this paradox may be explained, well within the framework of effect learning theory.

The response-unit hypothesis.—The first of these explanations has already been suggested. Various studies (7, 11) have shown that the reinforcing effects of reward apply not only to the immediately preceding behavior but also, to a diminishing extent, to the behavior which is temporally more remote from the advent of reward. According to this principle (the 'gradient of reinforcement'), we know that under the conditions of the present experiment it was possible for an animal to be somewhat reinforced for pressing the bar even when this behavior did not immediately produce food. Thus, because an animal was, for example, rewarded only after pressing the bar twice, we should not think of this as press-failure, press-reward, but rather as press-press-reward. It will be recalled that in this part of the experiment the animals were all trained, after the fourth day, according to some one definite pattern. The pattern for Group I was press-reward; for Group II it was press-press-reward; for Group III it was press-press-press-reward; and for Group IV it was press-press-press-press-reward. Therefore, instead of thinking of reinforcement as being restricted to the bar depression which occurred just before the reward, we should rather see the reinforcement as applying, in decreasing amount, to the other preceding depressions as well.

If we now reëxamine the results of this experiment in the light of the preceding discussion, we find that the discontinuous-reinforcement groups did *not* make more extinction 'responses,' defined in terms of the pattern, or unit, of bar depressions which was rewarded during acquisition, than did the continuous-reinforcement group. During extinction the animals in Group I repeated the response-unit which had previously been rewarded 128.0 times (128.0 ÷ 1), the animals in Group II repeated their response-unit 94.0 times (188.0 ÷ 2), those in Group III, 71.8 times (215.5 ÷ 3), and those in Group IV 68.1 times (272.3 ÷ 4). By thus using as a definition of 'habit,' or 'response,' the total unit of bar depressions which was rewarded in each of the four groups of animals during acquisition, we find that the apparent advantage of so-called intermittent reinforcement disappears (see solid line, Fig. 2). The hypothesis that rewarded repetitions of a habit (so defined) strengthen it is, therefore, still tenable.[7]

[7] Sears (27) has recently commented on the difficulty involved in precisely defining the term 'instrumental act' and has to a considerable extent anticipated the above analysis by pointing out that "all activity preceding the goal response appears to come under this heading and to undergo reinforcement. If a given *movement* must be performed twice before the goal is reached, then the *pair* of movements becomes the instrumental act to the goal response. If later, these acts are repeatedly performed without any reward occurring, it is to be expected that more *single* movements would occur during the extinction process if a *pair* of movements was the unit that had originally been reinforced than if a *single* movement had been the unit" (p. 84). Skinner (28) has similarly remarked: "This is fundamentally a problem in the definition of a unit of behavior. As a rather general statement it may be said that when a reinforcement depends upon the completion of a number of similar acts, the whole group tends to acquire the status of a single response, and the contribution of the reserve tends to be in terms of groups" (p. 300).

A subsidiary problem remains, however. Since we are proceeding on the assumption that every rewarded repetition of a habit (defined as the total response-unit, or instrumental act) strengthens it, and since each of the groups of *S*s in the present experiment made the same number of rewarded response-units during acquisition, each group might have been expected to make approximately the same number of response-units during extinction. But as Fig. 2 indicates, the average number of extinction response-units becomes progressively smaller as we go from Group I to Group IV, being 128.0, 94.0, 71.8, and 68.0, respectively. The *t*-value for the differences between the mean thus obtained for Group I and the means of the other three groups and between the mean for Group II and that for Group IV all show significance at the .05 level or better, thereby indicating that the downward slope of the solid line in Fig. 2 is statistically reliable and scientifically meaningful.

In a previous paper (**21**) we have shown that if a response requires more effort to perform, it will extinguish more quickly than if it requires less effort. Obviously, a response-unit consisting of four bar depressions requires more effort than a response-unit of only one depression. It is not surprising, therefore, that the former extinguishes more rapidly than the latter.

Equally plausible is an explanation based on the fact that even though all of the bar depressions involved in a response-unit of four depressions receive some reinforcement, they do not receive the same amount, since the one nearest the reward gets most, the one preceding somewhat less, the one preceding that one still less, and the original one in the series least of all.[8] On the other hand, if the response-unit consists of only one bar depression, it always receives the same, maximal amount of reinforcement (provided that the animal goes immediately to the trough and eats the food). Therefore, although the behavior involved in pressing the bar will be considerably more reinforced in Group IV than in Group I, it will not be reinforced *four times* as much. Consequently, when the behavior which occurs during extinction is analyzed in terms of response-units, Group I will necessarily rank highest, with Groups II, III, and IV following in order.

Whether one or both of the two mechanisms just described—greater effortfulness of multiple than of single bar-depressions, or the greater average but smaller total reinforcement received by single than by multiple depressions—were actually involved in producing

[8] Hull (**11**) has recently reported evidence that in order for a response on the part of a rat to be reinforced by a subsequent reward it must occur within 30 sec. of that event. In the present investigation, all of the bar depressions comprising a response-unit often occurred within a 30-sec. period (even in Group IV), but this was not always the case. (This statement is based on general observation; we have no systematic data.)

the downward slope of the solid line in Fig. 3, is a question which cannot be answered on the basis of our present data.

The discrimination hypothesis.—The second way of explaining the results obtained with intermittent reinforcement depends upon the concept of discrimination, or sign-learning. Let us, for the moment, disregard the immediately preceding analysis in terms of response-units. Viewing the continuous- and discontinuous-reinforcement procedures afresh, we see that the animals in Group IV, for example, had no means of differentiating so sharply between the acquisition situation and the extinction situation as did the animals in Group I. For the Group I animals, the act of pressing the bar had always 'meant' that food would follow, i.e., during acquisition the kinaesthetic and other stimuli which accompanied the performance of this movement became conditioned to the responses involved in turning from the bar to the trough and consuming the food. On the other hand, in the case of the Group IV animals the kinaesthetic and other stimuli accompanying bar depression were ambiguous, since they were sometimes followed by eating responses and sometimes by pressing the bar again, once again, and still again. Therefore, with the onset of extinction, when food did not follow the first, second, or even third depression of the bar, the situation was not yet any different, psychologically, from what it had been during acquisition. Only after at least four bar depressions had occurred and not been followed by food did the 'meaning' of the situation begin to change for the Group IV rats and then probably only very slowly. One would accordingly expect that the tendency to press the bar would wane much less quickly in the discontinuous- than in the continuous-reinforcement groups.

We suspect that this explanation is essentially what Humphreys had in mind in speaking of the greater difficulty which *Ss* who have been trained with discontinuous reinforcement have in changing from an 'expectation,' or 'hypothesis,' of reinforcement to one of non-reinforcement than do subjects who have been trained with continuous reinforcement. Our alternative statement merely attempts to make explicit some of the intermediate steps which were apparently implicit in Humphreys' reasoning.

The plausibility of the foregoing interpretation is increased by the observation, reported by Pavlov (**25**) and confirmed by others, that in learning experiments in which *Ss* are continuously reinforced, ex-extinguished, retrained, reëxtinguished, etc., extinction and re-learning both take place with increasing speed. This finding suggests that the receiving of the reward has psychological implications over and beyond the automatic action of the Law of Effect. It suggests that

the absence of reward, on the first trial of successive extinctions, may come to signify to the *S* that further response is, at least for the time being, futile, i.e., the stimuli associated with the absence of reward become conditioned to resting instead of working, just as the presence of reward, on the first trial during successive re-learning periods, may come to signify to the *S* that the opposite situation now prevails.[9] Thus, in an experiment in which hunger is the principal drive, the *S* learns to discriminate sharply between two total patterns: hunger-(and-fatigue)-plus-stimuli-associated-with-getting-food and hunger-(and-fatigue)-plus-stimuli-associated-with-*not*-getting-food. To the first configuration the *S* learns to make food-seeking responses; to the second configuration the *S* learns to make resting responses.[10]

It is obvious that if reinforcement is discontinuous during acquisition, acquisition will be harder to discriminate from extinction (ergo, greater 'resistance to extinction') than if reinforcement has been continuous. And not only would we expect to encounter this type of confusion on the first transition from acquisition to extinction; if, as

[9] Cook and Harris (3), Cole (4), and Mowrer (17) have shown that it is possible to produce almost instantaneous 'extinction' and 're-learning' of involuntary conditioned responses in human *S*s by means of signals or verbal instruction to the *S*s indicating changes in experimental conditions. All these investigators used avoidance rather than appetitive conditioning, which probably means that they were simply manipulating the secondary motive of fear symbolically instead of requiring their *S*s to build up their own inductive generalizations, or 'hypotheses.' *Cf.* Krechevsky (16) and Brunswik (2) for a discussion of this problem as it relates to the behavior of rats.

[10] In acknowledging the 'sign-function' of success and failure, one does not need to repudiate the Law of Effect, as Tolman, Ball, and Brentnall (29), for example, suggest. It would appear that we are dealing here with a mechanism which often supplements and on occasion even supersedes simple law-of-effect learning, but this is no proof that the latter is not a perfectly valid phenomenon. Our guess is that in those situations in which one non-rewarded repetition of a response (or a signal of some kind—see footnote 9) causes a sudden cessation ('extinction') of response what happens is that a strong *emotion* (anticipation of many non-rewarded, futile performances of the response) is aroused which serves to counteract the habit strength, or response tendency, built up by the preceding rewarded repetitions of the response. Likewise, a single rewarded repetition of a response (or a signal from the experimenter—see footnote 9) may likewise serve as a kind of all-clear sign which removes the depressing effect of many antecedent non-rewarded repetitions of a response (or of an earlier signal from the experimenter) and thus releases the tendency to make the response in question. Or it may be that a single success or some other sign that reward will now follow if responses are resumed may serve, not merely to banish 'discouragement' but also to arouse an emotion of 'hope' which suddenly energizes the response tendency. Thus a single success or failure (or a signal or word with the same implications) may cause an abrupt change in behavior of the kind that has often been accepted as proof of 'insight,' 'change in hypothesis,' or 'alteration in cognitive structure' and as disproof of the Law of Effect. In our judgment this is not a real antithesis and should not cause the cleavage which exists today between those learning theorists who stress the role of effect and those who stress so-called cognition. What we are suggesting is that *emotion* is the key to the controversy and may enable us to build up a more unified structure of behavior theory, from the lower to the higher levels of complexity, than now exists. The present hypothesis would seem to have especially promising applications to discrimination learning and in resolving such paradoxes as Hull (11) has recently pointed out in connection with 'patterning.' It may also be useful in more correctly interpreting 'latent learning.'

in Pavlov's experiment, we passed a number of times from the one condition to the other, we should expect less rapid discriminative learning where acquisition involved discontinuous reinforcement than where it involved continuous reinforcement.

This type of analysis turns out to have an unsuspected advantage: it suggests why it is that in previous experiments on continuous vs. discontinuous reinforcement, there has been the confounding, already noted, of the continuity-discontinuity factor with the regularity-irregularity factor. The research in this field was started by Humphreys, using college students as subjects.[11] Knowing that such *Ss* can *count*, he may have anticipated that the effects of discontinuity could be demonstrated only if he made the reinforcement procedure not only discontinuous, but also irregular, since by counting a human being can detect a change from an acquisition to an extinction situation quite as accurately if the reinforcement has been discontinuous as if it has been continuous, provided the reinforcement has been regular. In other words, for a human being who, during discontinuous but regular reinforcement, makes each unrewarded trial distinctive by performing the 'pure stimulus act' (Hull, **9**) of saying, 'One,' 'Two,' 'Three,' etc., failure of reward to follow, for example, a fourth trial after three non-rewarded ones can be just as significant as the occurrence of a single non-rewarded trial following continuous reinforcement.

As Reiss (**26**) has pointed out, the rudimentary function of counting is to make possible the same sharpness of discrimination between 10 and 11, or 100 and 101 otherwise identical, successive events (reactions) as between 0 and 1 or 1 and 2 such events. With such precision of discrimination possible by means of counting, it is questionable whether the Humphreys' effect can be demonstrated in educated human *Ss* with discontinuous but regular reinforcement.

Regular vs. Irregular Reinforcement

The foregoing discussion provides a convenient setting for the description of the results obtained from our Group V animals, which, it will be recalled, received reinforcement both intermittently and randomly, but on the average after every two and one-half depressions

[11] Skinner's work (**28**) on discontinuity of reinforcement with rats actually antedates that of Humphreys, but it was only with the publication of Humphreys' findings that general interest in this problem was stimulated. Brogden (**1**), Cole (**4**), and Brunswik (**2**) have also published data on the effects of intermittent reinforcement, showing that, in general, acquisition of a given habit occurs more slowly with intermittent than with continuous reinforcement (*cf.* also Finger, **5**). Of these three writers Cole alone was interested in the possible effects of intermittent reinforcement on resistance to extinction and he only incidentally. Insofar as his results can be compared to those of Humphreys and Finger they are, however, confirmatory.

of the bar. These animals depressed the bar a mean number of 189.2 times during the three days of extinction. The point corresponding to this value has been indicated in Fig. 3. Dividing 189.2 by 2.5 (the figure comparable to the response-units of the other four groups), we find that the Group-V animals made 75.5 repetitions of their hypothetical average response-unit. The point corresponding to this value has likewise been indicated in Fig. 3.

From these results it is evident that the greater resistance to extinction obtained in this experiment following discontinuous (Groups II, III, and IV) than continuous (Group I) reinforcement is probably in no way contingent upon the factor of irregularity of reinforcement. In fact, as a glance at Fig. 3 will indicate, our Group V animals depressed the bar actually somewhat less frequently during extinction than would have been expected from interpolation on the curve representing the performances of the regular reinforcement groups (I, II, III, and IV).

Unfortunately, there is no way of determining whether the obtained downward deviation of the Group V results is due to the factor of irregularity or to factors which the experimenters did not attempt to control, i.e., 'chance.' If the Group V animals had been irregularly rewarded after pushing the bar either two or three times, on the average, instead of 2.5 times, their extinction scores would then have been directly comparable, statistically, to those for the Group-II or the Group-III animals.

That this procedure was not followed is now a patent defect in the design of the experiment. However, from a common-sense standpoint, our results indicate clearly enough that, for rats at least, it is discontinuity of reinforcement, rather than irregularity, that accounts for the greater resistance to extinction resulting from a training procedure involving both discontinuity and irregularity of reinforcement. Therefore, the confounding of the discontinuity and irregularity factors which has been noted in previous studies is probably unimportant and was even necessary if this phenomenon was to be demonstrated in human Ss, due to the latter's tendency to count and thereby render discontinuous and continuous reinforcement much more nearly equivalent, psychologically, than they otherwise would be.

Other Implications

As we have seen, there are two about equally plausible ways in which the Humphreys effect can be accounted for: by the 'response-unit hypothesis' and by the 'discrimination hypothesis.' Naturally the question arises as to which of these is the more valid, and experiments can probably be devised (though not easily) which would

throw light on this problem.[12] However, in situations of this kind we fall perhaps too easily into the practice of taking an either-or attitude, of assuming that where two or more mechanisms may account for a given result one must ultimately prove significant and the other not. Actually we know that in all but the most excessively simplified types of psychological situations the observed outcome is nearly always 'over-determined,' i.e., is a product of multiple processes acting, and interacting, simultaneously. It may well be, therefore, that *both* of the explanations advanced above to account for the Humphreys effect are valid.

One other implication of research on continuous vs. discontinuous reinforcement in relation to 'resistance to extinction' should be noted. In an experiment reported by Whiting and Mowrer (**30**), it has been discovered that when rats, previously punished for taking the shorter route to the goal on a D-shaped maze, found that the punishment (an electric shock) was no longer in force, they precipitously reverted to this route. This behavior was contrasted with the behavior of human beings, who often show life-long tendencies to refrain from doing certain things they have been taught to regard as 'wrong' even though the external sanctions originally used to set up such attitudes have long since disappeared.[13]

This difference between the behavior of rats and human beings raises the question as to what it is about the education, or 'socialization,' of the latter that creates conscience, character. The psychology of conscience is undoubtedly exceedingly complicated, but one consideration seems obvious: in the experiment reported by Whiting and Mowrer, the rats discovered that the punishment always occurred at a certain point along the shorter route to the goal. If they got past this point, they were safe and had no reason to feel 'guilty.' In

[12] The response-unit and the discrimination hypothesis are designed to account for the greater resistance to extinction (in the sense of more "work" done) following discontinuous than continuous reinforcement. They make no assumption concerning the effect of regular vs. irregular reinforcement. That the findings of this investigation indicate that, in rats at least, the factor of regularity-irregularity is of little or no importance in this connection *may* have implications for the relative validity of these two hypotheses, but if so these implications are so subtle as to be obscure to the present writers.

[13] Whether the originally (usual parental) sanctions become 'introjected' and form a permanent part of the personality ('super-ego') needing no further reinforcement from the external world or are subtly strengthened periodically by other (community) sanctions is one of the important and difficult psychological problems of our time. It is perhaps the central point of issue between Freudian psychoanalysis, with its emphasis upon the past (childhood), and the followers of Karen Horney and members of the Washington ('inter-personal relations') School of Psychiatry, for whom the current life situation of the individual is of greatest importance. This problem is also related to the question as to whether emotions (attitudes) are 'autonomous' (Allport) or require at least occasional reinforcement, like any other learned response. Professor F. J. Shaw (personal correspondence) has remarked, "We still face the problem of the apparent resistance of the anxiety state itself to extinction. This strikes me as a rather difficult problem for reinforcement theory whereas Guthrie's theory might have less difficulty with this particular problem."

learning to run the 'social maze' which parents and others set for them, children, on the other hand, discover at an early age that punishments, even though indefinitely delayed, may still be forthcoming; and certainly parents are likely to foster the impression on the part of their offspring that their 'sins' must ultimately be found out and that retribution will surely follow.

Thus it is that at least a part of that kind of conscience which is based on fear is probably set up. But if conscience is an institution whereby the individual punishes himself (i.e., his conscience 'hurts' him) for acts, contemplated as well as actually performed, which are likely to bring external disapproval, so is there a comparable mechanism, based on hope and faith, which makes for perseverance in the face of adversity, in a word, 'morale.' Very small children display both the moral and the morale aspects of conscience in only incipient amounts, and much of their 'character training' consists of attempts to bring about development along these lines. As part of such a program they are commonly encouraged to keep plugging away if they do not immediately reach their (legitimate) goals and are admonished, when confronted by failure, to 'try, try again.'

Even in the course of ordinary events, children learn to carry out long sequences of action where the reward is so remote that in the beginning it can sustain endeavor for only a short time. Thus, a small child, though capable of endless activity at home, will quickly report being 'tired' if taken for a walk. But it is almost certainly not so much fatigue as it is 'extinction,' or 'loss of interest,' that is responsible for his lagging steps. He puts one foot forward, then the other, then the one, then the other, etc., etc., and nothing happens. Only gradually, with encouragement from his elders and periodic indulgences—a look at the train here, an ice cream or candy there—does the child get so he can, at the age of five or six, set out on a long walk with only a single destination or objective in mind. His 'resistance to extinction' has increased, in a manner probably not fundamentally different from that used in producing the Humphreys effect. By requiring rats to push a bar at first only a few times for a piece of food, then a larger and a larger number of times, Skinner (**28**) was able to get them eventually to make 192 responses for a single piece of food, and there was no indication that this number necessarily represented the limit of which the rat was capable. Did these rats have 'morale'? They certainly had 'hope,' 'faith.'

Somewhere between fear and faith—the pivotal concepts of most religions—lies the truth about conscience. Any improvement in our understanding of how they work is a contribution to the psychology of character, and character defects. Here, it would seem, is a most promising point of articulation between scientific learning theory and

education, ethics, psychotherapy, and other practical problems of both a social and individual nature.

Summary

In this paper an attempt is made to show that the Humphreys paradox—that habitual responses which are rewarded only intermittently are apparently stronger than those which are rewarded after each occurrence—does not invalidate the Law of Effect. Experimental results are reported which confirm Humphreys' empirical findings but which can be explained either by the 'response-unit hypothesis' or by the 'discrimination hypothesis,' both of which are consistent with the view that reward is the fundamental and crucial determinant of habit strength.

The response-unit hypothesis defines a 'response,' not as a single, isolated movement, but as the totality of behavior which leads to a given goal. For example, the fact that a rat has to press a bar three times in order to get a pellet of food does not mean that three separate acts, or 'responses,' have been performed; the sequence of three bar depressions is here conceived as a unitary, integrated, instrumental performance. When 'response' is thus re-defined in terms of the whole pattern of behavior which proves effective in producing reward during acquisition, the Humphreys paradox disappears. When 'response' is thus re-defined, it is found that intermittent reinforcement, far from producing greater 'habit strength,' actually produces reliably less than does continuous reinforcement.

But there is another equally plausible approach to this problem. If, during acquisition, a response (conceived as a more or less isolated movement) occurs repeatedly but is rewarded only now and then, the transition from acquisition to extinction will not be discriminated as sharply as if acquisition has involved reward for each and every response. With 'faith' thus established that failure will ultimately be followed by success, 'discouragement' is slower to set in (ergo, greater 'resistance to extinction') when there is a change in objective conditions from acquisition (occasional reward) to extinction (no reward whatever).

This second interpretation involves the introduction of *intervening variables* ('faith,' 'discouragement'), factors which disrupt the direct, one-to-one relationship observed in the simplest kinds of situations (and subjects) between reward and behavior. Here we are approaching the realm of 'ego psychology,' but it should not be supposed that there is any discontinuity between the fundamental laws of learning that can be so easily demonstrated at the lower levels of behavior and the generalizations that seem to hold best at the higher levels. The

intervening variables which operate at the higher levels and which complicate the simple relationships observed at the lower levels between reward and the behavior-changes which we call learning are themselves products of former learning—i.e., 'emotions,' 'secondary drives,' 'expectations,' 'attitudes,' 'sentiments,' 'conscience.' [14]

Here we are dealing with a distinction similar to the one which Murray (**24**) has made between *viscerogenetic needs* and *psychogenic needs*. And when there are psychogenic needs (secondary drives) operating within an organism, we do not expect the viscerogenetic needs (primary drives) to produce the same results as when psychogenetic needs are absent. When these two types of 'needs,' or drives, come into opposition, there will necessarily be *conflict* (sometimes 'character'); and we know that the outcome is rarely a matter of simple algebraic summation (**23**). The various 'dynamisms' (Freud), or 'self-systems' (Sullivan), go into action, with the endlessly complicated results which we see clinically (as well as in the 'experimental neuroses' of animals) and even in ordinary, 'normal' human behavior.

(Manuscript received February 1, 1945)

References

1. Brogden, W. J. The effect of frequency of reinforcement upon the level of conditioning. *J. exp. Psychol.*, 1939, **24**, 419–431.
2. Brunswik, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25**, 175–197.
3. Cook, S. W., & Harris, R. E. The verbal conditioning of the galvanic skin reflex. *J. exp. Psychol.*, 1937, **21**, 202–210.
4. Cole, L. E. A comparison of the factors of practice and knowledge of experimental procedure in conditioning the eyelid response of human subjects. *J. gen. Psychol.*, 1939, **20**, 349–373.
5. Finger, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
6. Finger, F. W. Retention and subsequent extinction of a simple running response following varying conditions of reinforcement. *J. exp. Psychol.*, 1942, **31**, 120–134.
7. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century Co., 1940.
8. Hovland, C. I. 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. nat. Acad. Sci.* (Washington), 1936, **22**, 430–433.
9. Hull, C. L. Knowledge and purpose as habit mechanisms. *Psychol. Rev.*, 1930, **37**, 511–525.
10. Hull, C. L. The excitatory-potential hypothesis and behavior theory. On file in Institute of Human Relations, Yale Univ., 1938.
11. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., 1943.
12. Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.

[14] We do not overlook the fact that 'thoughts' and 'ideas' are also important intervening variables, but the intra-psychic use of symbols is so intimately tied up with emotions that in stressing the latter we are necessarily subsuming the former. It is a common form of human vanity to make thoughts a 'higher' and emotions a 'lower' mental process, but from a dynamic standpoint, the emotions are both genetically and energetically basic. Emotions supply the rudiments of 'reason' and remain its driving force no matter how refined or elaborated it becomes (**23**).

13. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.
14. Humphreys, L. G. Extinction of conditioned psycho-galvanic responses following two conditions of reinforcement. *J. exp. Psychol.*, 1940, **27**, 71–76.
15. Humphreys, L. G. The strength of a Thorndikian response as a function of the number of practice trials. *Psychol. Bull.*, 1940, **37**, 571.
16. Krechevsky, I. Antagonistic verbal discrimination habits in the white rat. *J. comp. Psychol.*, 1932, **14**, 263–277.
17. Mowrer, O. H. Preparatory set (expectancy)—a determinant in motivation and learning. *Psychol. Rev.*, 1938, **45**, 62–91.
18. Mowrer, O. H. Motivation and learning in relation to the National Emergency. *Psychol. Bull.*, 1941, **38**, 421–431.
19. Mowrer, O. H., & Lamoreaux, R. R. Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.*, 1942, **54**, No. 5, 34 pp.
20. Mowrer, O. H. A cumulative graphic work-recorder. *J. exp. Psychol.*, 1943, **33**, 159–163.
21. Mowrer, O. H., & Jones, H. M. Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, 1943, **33**, 369–386.
22. Mowrer, O. H., & Lamoreaux, R. R. Fear as an intervening variable in avoidance conditioning. *J. comp. Psychol.* (in press).
23. Mowrer, O. H., & Ullman, A. D. Time as a determinant in integrative learning. *Psychol. Rev.* (in press).
24. Murray, H. A. *Explorations in personality.* London: Oxford Univ. Press, 1938.
25. Pavlov, I. P. *Conditioned reflexes.* (Trans. and ed. by G. V. Anrep.) London: Oxford Univ. Press, 1927.
26. Riess, A. An analysis of children's number responses. *Harvard educ. Rev.*, 1943, **13**, 149–162.
27. Sears, R. R. *Survey of objective studies of psychoanalytic concepts.* New York: Social Science Research Council, 1943.
28. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century Co., 1938.
29. Tolman, E. C., Hall, C. S., & Brentnall, E. P. A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption. *J. exp. Psychol.*, 1932, **15**, 601–614.
30. Whiting, J. W. M., & Mowrer, O. H. Habit progression and regression—a laboratory study of some factors relevant to human socialization. *J. comp. Psychol.*, 1943, **36**, 229–253.

EFFECT OF VARIOUS CONCOMITANT ACTIVITIES UPON THE MAINTENANCE OF A SET

BY DOROTHY RETHLINGSHAFER

Woman's College, University of North Carolina

The analytical unit of the stream of behavior appears today to be an act, or task, or activity. Watson's concatenations of reflexes have been abandoned in favor of the larger units, but the interrelationship of these behavior units has not been investigated.

We know that the behavior stream may be carrying more than one activity. We know that the higher organisms may be set to respond to an absent stimulus, and that during the maintenance of that set, they may be engaged in various other activities. In order to investigate the effect of concomitant activities upon the maintenance of a set, the *S*s were (1) directed to react as quickly as possible to an auditory stimulus—a buzzer—whenever it appeared, and (2) along with this intermittently overt task, to engage in a variety of other activities. Two of these activities were learning problems (maze and memorizing nonsense syllables), one involved insight, and two were puzzles of varying difficulty.

The problem was to determine how the *S*s might vary in the continuous, though only intermittently overt, task of stopping the buzzer whenever it sounded while engaged in various other activities. The problem was limited to determining only the changes that would take place, if any, in the reaction time to the buzzer during the course of the other activities. All measurement of the interrelation of the behavior units was in terms of changes in the reaction time to the buzzer as taken during different stages of the performance of the concomitant activities.

Method, Apparatus, and Subjects

The reaction time was measured by a Dunlap chronoscope so wired in circuit with a buzzer that when *E* closed his key, the buzzer sounded until *S* reacted by closing his key and holding it down. This broke the buzzer and stimulus coil circuit and activated the response coil so that the reaction time could be read from the dial. It should be noted that the response coil would not work unless *E*'s key was down. *S* was instructed not only to stop the buzzer as quickly as possible by pressing his key, but to hold his key down until he heard *E* say 'O K.' If *S* failed to follow directions, the error was quickly revealed by the starting dial.

Woodworth (7) has pointed out the possible source of error in the clutch of the Dunlap chronoscope, and an examination of our results shows that the mean of the 10 records obtained from the 43 *S*s used with the ready signal is greater than the average results for an auditory stimulus. However, only two practice trials were given, though there may have also been a con-

stant error in the chronoscope. Variability as indicated by the sigmas obtained was higher than in other studies of reaction time. The reliability of the records obtained by correlating the first half and last half was .92 when the Spearman-Brown prophecy formula was used.

(1) The first reaction time records to the buzzer were obtained with a *ready signal* and a slightly varying foreperiod of approximately one second. Only two practice trials were given, and these were not included in the data. No attempt was made to differentiate between sensory and muscular preparation, though the directions of stopping the buzzer as quickly as possible may have led to listening for the buzzer and hence to sensorial reactions.

(2) The *S*s were next told that *no ready signal* would be given but that they were to stop the buzzer as quickly as possible whenever it sounded. *E* sat opposite *S* and his key was behind a screen. Two electric fans in the room masked any noise of the apparatus and, to some extent, noises outside the room. The buzzer, which was well above the auditory limen of any of the *S*s, was sounded at the following irregular intervals: 10, 20, 45, 45, 35, 25, 50, 30, 15, 45, and 30 sec., so that 11 reaction time records were obtained during five min. and fifty sec.

(3) Then the *S*s learned eight nonsense syllables taken from the 12 percent level of Hull's list (5), and at the same time they were told to stop the buzzer as quickly as possible. During all the activities the *S* sat with the index finger of his right hand resting lightly on his reaction time key. In order to remind the *S* that he must hold his finger on the key, his right arm was strapped very loosely in a comfortable position on the table.

The nonsense syllables were presented by using the Missouri memory apparatus with the exposure time at two sec. Only one reaction time to the buzzer was obtained during the presentation of the list of eight syllables. The buzzer sounded by following a standard system for each *S*, i.e., immediately after the first syllable was pronounced, then in the next presentation after the sixth syllable, etc. The criterion of learning was four perfect recitations.

(4) The second activity used with the reacting-to-the-buzzer task was learning a stylus maze. The *S*s were told that it was not a speed test.[1] As the *S* took the stylus in his left hand, *E* repeated the directions for the maintenance of the set to the buzzer: "Remember you are to stop the buzzer as quickly as possible whenever you hear it." The *S*'s right hand was, of course, ready at the reaction time key. One reaction time was obtained during each trial, at a standard place in the maze for each *S*. The buzzer was sounded in the first trial after the *S* had passed the first blind, then in trial two after blind five, in trial three after blind six, in trial four after blind four, etc. The criterion of learning was four perfect trials.

(5) Following the maze, the *S*s were given a task described as an insight problem by Pickford (6). Twenty-seven straight line figures were drawn on cards. Fourteen of these figures had a right angle within the inner area. These were mixed with the non-right angle figures. The *S*s were asked to examine the figures one at a time and when they had discovered the characteristic common to certain cards to report their discovery to the experimenter. Again they were reminded before they started the insight problem to stop the buzzer as quickly as possible. Most of the *S*s stated the principle before completing three trials. In case the *S* had not seen the correct characteristic, the activity was discontinued after three trials.

Five reaction records to the buzzer were obtained from each *S* during the first trial, and four thereafter. Just as the *S* was starting the task of discovering the common element in the cards, one reaction time was obtained, then again the buzzer was sounded after cards 4, 10, 18, and 25.

If an *S* had insight, then he was further asked to continue through the 27 cards indicating whether each card had the characteristic of right angles, and where he had first 'got the idea.' Four more reaction times to the buzzer were obtained during this 'after-insight' activity. Insight into the problem was assumed on the basis of the following: (*a*) verbalization of a general principle, and (*b*) disappearance of errors after the statement of the principle.

(6) The next activity during which the *S*s carried on their reaction time task was a 17-piece wooden puzzle, painted white. The directions were: "These pieces will fit into this space. (Rectangle indicated) I want you to put them in. You are also to stop the buzzer as quickly as possible."

Reaction time records were obtained as follows: after the directions but before any successful moves on the part of the *S*s; after three blocks were successfully fitted together; then after 6, 9, 12, and 15 blocks were fitted together; and finally just after completion. Since some of the *S*s

[1] Daniel (1) has indicated that there is an effect upon the tension level of an individual depending upon his purpose, the elimination of errors or speed.

pulled away their fingers from the reaction time key as soon as they completed the puzzle, records of five *Ss* were not used in the 'after completion' mean.

(7) A second puzzle, more difficult than the first, was presented, and again the reaction time to the intermittently sounding buzzer was also demanded. The directions were similar to the above. Five records were obtained from each *S*: after directions but before starting the puzzle; when approximately three blocks had been fitted correctly; when approximately eight blocks had been fitted correctly; when eleven blocks had been fitted correctly; and after completion. With certain *Ss* records were also obtained at three-min. intervals when they came to a difficult place in the puzzle. Seventeen individuals found one place quite difficult, and 31 records of reacting to the buzzer were obtained while they were not making any successful moves.

Subjects

Forty-three college students of the University of North Carolina served as *Ss*. Though all the *Ss* were tested for their reaction to the sounding buzzer during the various activities described above, a few records were discarded, due to the fact that the *S*'s hand was not always in a correct position at his reaction key.

Results

Reaction time to the buzzer with signal and without signal.—In the usual reaction time experiments, the ready signal serves to reinstate the *S*'s set to react as quickly as possible whenever the stimulus reappears to which he is directed. If, however, the reaction time is obtained to this same stimulus without a ready signal, the time increases even though no experimentally introduced 'extra' tasks are given the subject. As can be seen in Fig. 1, there is a consistent difference between these two conditions—with ready signal and without ready signal. This finding is, of course, not new, but gives us a base in this study from which comparisons can be made with the reaction time obtained when concomitant activities are experimentally introduced into the stream of behavior.

Maintenance of a set to react to the buzzer when the Ss are also engaged in learning problems.—Two learning problems were used as concomitant activities with the reaction time task. As can be seen in Fig. 1, there is at first long reaction time and then a decrease as the learning activities become more automatic. This is true for both the learning of the stylus maze as well as the memorizing of the nonsense syllables. In the first two trials of both learning problems, when 48 percent of the errors were made, the mean reaction time was 958 sigma for the stylus maze and 918 sigma for the nonsense syllables. In the remaining error trials the reaction time decreased approximately the same amount during both activities. Interestingly, however, there was also a continued drop in the reaction time to the buzzer during the remaining four errorless trials. The efficiency with which the *Ss* maintained their set to react to the buzzer continued to improve even after the *Ss* had stopped making errors in their performance of the concomitant activities. The critical ratio between

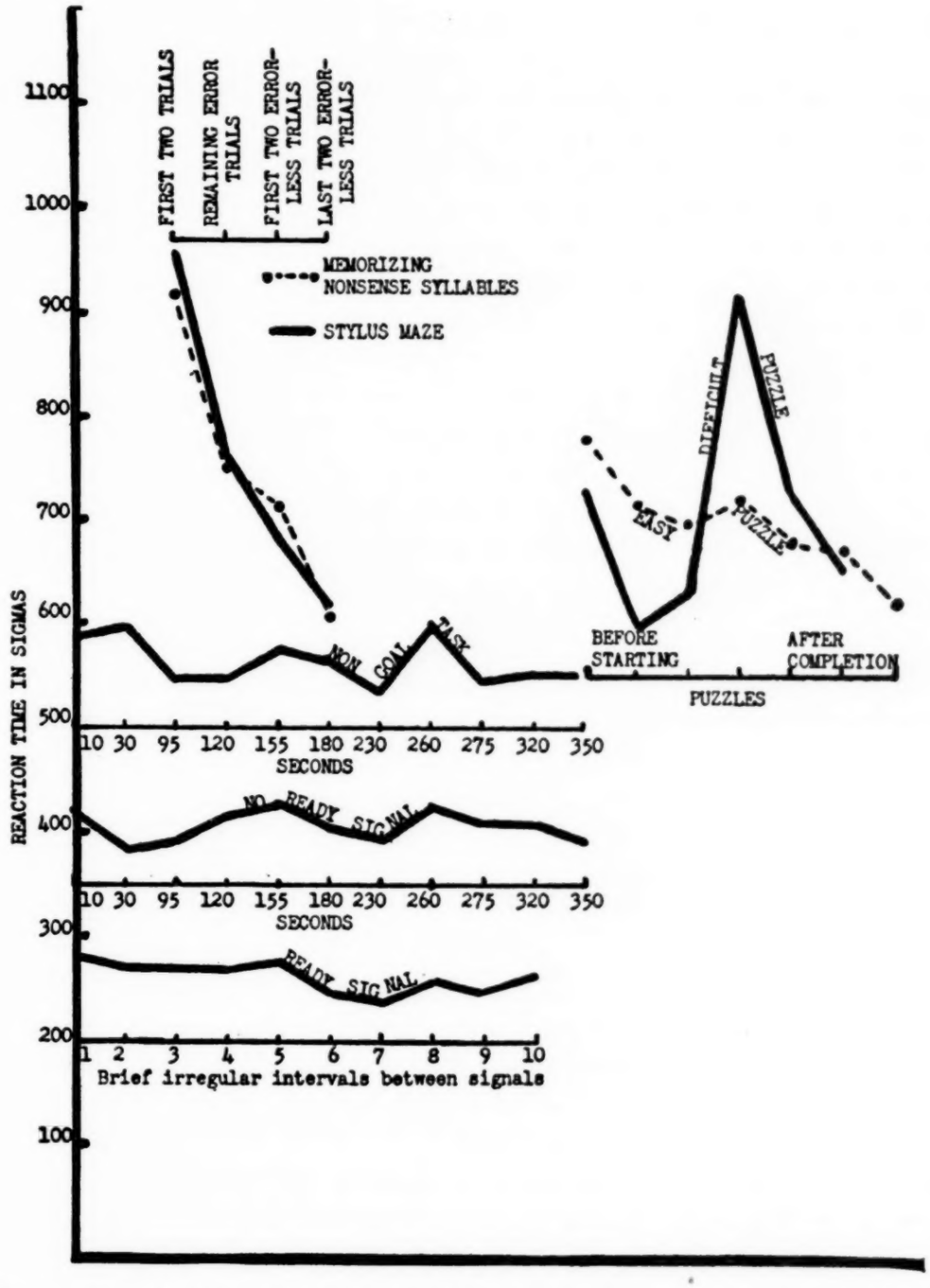


FIG. 1. Reaction times obtained in stopping a buzzer when *Ss* were also engaged in various other activities.

the reaction time taken during the first two trials and the last two trials is above three for both the learning problems.

Maintenance of a set while the Ss were engaged in an insight problem.—Records in reaction time to the buzzer were obtained while *Ss* were engaged in identifying the principle which distinguished certain polygons from others. While the *Ss* were seeking to verbalize the common characteristic—possession of a right angle—records of reac-

tion time to the buzzer were obtained at predetermined intervals. Four *Ss* were never able to verbalize the principle, i.e., showed no insight. In Table I is given the mean reaction time to the buzzer of the 35 *Ss* who succeeded and of the four who failed.

The successful *Ss* in the insight problem significantly decreased in reaction time following their success. The critical ratio of the difference between the means, before-insight and after-insight, was 5.3. Though this decrease was significant, the average sigma of the after-insight records was greater than obtained finally during the last two errorless trials of the learning problems. After-insight records were obtained while the *Ss* were being tested for their ability to use the principle that they had just verbalized. Though this new activity was not difficult, it may have interfered with the maintenance of the set to react to the buzzer. For this reason, 'success' in the insight problem may not have brought as great a decrease in reaction time as did 'success' in the other learning problems.

TABLE I

REACTION TIME TO A BUZZER WHILE SOLVING A PROBLEM CALLING FOR INSIGHT

	Mean	S.D.	*N*	Number of Records
Before Insight.....	897	289	35	143
After Insight.....	731	236	35	133
No Insight.....	660	189	4	51

The four *Ss* who were unsuccessful in gaining insight were slightly superior in maintaining the set to react to the buzzer and their variability was less. Were these four *Ss* unsuccessful in the insight problem because they were maintaining the buzzer-set at a level which prevented success in the other task?

Difference in the maintenance of the set for the buzzer during the solving of a 'difficult' puzzle as contrasted to an 'easy' puzzle.—Two puzzles, each consisting of wooden blocks to fill designated space, were used with the reaction time task, though, of course, presented separately. The criterion of difficulty which differentiated the puzzles was simply time for completion. It was also known from previous work that the progress of some *Ss* would be temporarily halted at a particular place in the difficult puzzle. When a certain relationship necessary for further placing of the blocks was understood, then the *Ss* could almost immediately complete the puzzle. Seventeen of the 43 *Ss* were stopped by this particular difficulty, and 31 records of reaction time to the buzzer were obtained while they were seeking by trial and error for the correct solution. As can be

seen in Fig. 1, the reaction time increased significantly at this point in the solution of the difficult puzzle.

Six records were obtained from each *S* as he worked upon the difficult puzzle. The first record was the reaction time of the *S* to the buzzer, sounding after the directions had been given but before the *S* had started to place the blocks in the designated space. The mean performance at this point was slightly higher than when approximately three blocks were fitted successfully together. As Fig. 1 illustrates, however, the reaction time to the buzzer increased as the puzzle became more difficult, reaching a mean performance of 913 sigma. The difference between this sigma and that of 595 at the third successful block is significant with a critical ratio of 6.2. Another record was obtained when the *S* had almost finished the puzzle and the mean performance is back to 729 sigma. A further drop, though not so great, was recorded when the *S*s reacted to the buzzer immediately after completion.

A similar picture is presented in Fig. 1 for the reaction time to the buzzer as the *S*s work upon the simpler puzzle. As the subjects 'get into' the puzzle, a slight increase in reaction time is recorded, but there is a more gradual and continuous fall in the reaction time, as shown in Fig. 1, by contrasting the records taken before starting with those after completion. The critical ratio of the difference between the first records and those obtained after completion is 2.5.

During both puzzles an improvement is shown as the *S*s progress from the introduction into the concomitant tasks to the final completion. The buzzer-set is most seriously affected when the accompanying task becomes difficult.

Efficiency of the buzzer-set while the Ss are also engaged in a non-goal task.—In order to determine how the reaction time task would be affected by the concomitant carrying of a non-goal activity, the *S*s were directed to drop a marble through a hole in a lid of a three-sided box, pick up the marble, drop it again, and so on until told to stop. The reaction time to the buzzer was obtained at the same intervals used with the no-signal series when there was no experimentally introduced task present. Of course, the *S*s were 'free' to engage to a limited extent in their 'own' activities during the no-ready signal series, and also, no doubt, were able to introduce 'goals' into their 'non-goal' task. A comparison of the reaction times obtained, however, under these two different conditions shows that the buzzer-set was significantly affected by even a simple non-goal task. The critical ratio of the smallest difference is significant between the non-ready signal records and the non-goal task.

Examining Fig. 1, we find no significant trend in the rise and fall of the reaction time as the *S*s engaged in the non-goal task. This is in

marked contrast to the drop recorded during the learning problems and during the solving of the puzzles. Since there is no goal to reach, no special barrier or difficulty to overcome, no possibility of having insight or of changing from errors to success, the buzzer-set showed no marked changes during this concomitant activity. At the eighth test for its efficiency, the records indicate that it is maintained even slightly less ably than at the beginning of the non-goal task. The mean is 597 sigma as contrasted to the mean of 589 at the beginning of the marble-dropping.

Discussion

When there is added in the stream of behavior any other concomitant activity to the buzzer-set, its efficiency is decreased. And when the accompanying activity is a goal task, the reaction time is more seriously affected. The completion of the concomitant activity, either in terms of solution or of errorless performance, brings an improved set for responding quickly to the buzzer, but when any barrier is present in the extra activity, the reaction time is longer. And it seems to follow from the data on the difficult puzzle that the stronger the barrier to completion of the accompanying task, the more the set for responding to the buzzer is decreased in efficiency.

It is felt, of course, that the above results are dependent upon the conditions of presentation of the two tasks as described in the procedure. The attempt was made to emphasize equally the reaction time task and the concomitant activities. Hence the increases and decreases in the reaction time are the result of greater or lesser 'demands' from the extra activities upon the organism. It would have been possible, of course, to have increased the 'significance' of responding to the buzzer so that all the *S*s would have maintained the buzzer-set at a level of efficiency higher than herein shown. It is the guess of the writer that under a greater motivation for the buzzer-set, the differences between the means would have been lowered, as well as the general level of all the means, but that the same trends would have appeared as shown in Fig. 1.

Raising the importance of the buzzer-task in the hierarchy of the behavior stream would be possible and no doubt would affect the performance of any concomitant activities. No attempt was made to measure the interaction of activities in this manner.

It should be noted that the set to react to the buzzer was never lost. Though some *S*s became so 'absorbed' in the concomitant activity that the buzzer sounded over 1000 sigma before the *S*s reacted by pushing down the key, nevertheless no one had to be redirected to the buzzer. Of course at the beginning of each new

activity, the set for the buzzer was reëstablished by repeating the directions: "Remember you are also to stop the buzzer as quickly as possible." Thus, the intermittently overt task maintained itself though at varying levels of efficiency during the extra activities imposed upon it.

Interpretation in Terms of Shifting Tension

A possible interpretation of the changes shown in Fig. 1 is in terms of shifting tension as the concomitant activities are added to the behavior stream. It may be assumed that optimal reaction time depends upon optimal tension in the reacting arm muscles. The experimental data herein presented do not determine whether the tension load of the reacting arm for the buzzer task was *below* or *above* the optimal when the longer reaction time records were obtained. Both possibilities are presented below.

It could be assumed that the tension load of the arm and hand used in stopping the buzzer decreased *below* the optimal when the other activities demanded that general body tension be directed at their solution. At certain places in the course of these concomitant goal activities, i.e., at the beginning of the learning problems and at strong barriers in the puzzles, there has been some evidence (4) of an increased tension load. At these same places we found longer reaction time in our task of stopping the buzzer. Does this indicate a shifting of tension away from the right arm and hand necessary for the reaction time task? Since we found with success or completion in the concomitant activities a shorter reaction time in the task of stopping the buzzer, we may assume that the tension then had shifted back to the reacting arm and thus was nearer the optimal for a reaction time response.

On the other hand it can be assumed that a muscular condition in any part of the body may influence the excitation level of any other muscle by way of the proprioceptive pathways. Why, therefore, should the tension load decrease in the arm set to react to the buzzer, when, as at the beginning of an activity, body tension was generally increasing? Were we wrong in assuming that the longer reaction times were due to a level of arm tension below the optimal? Freeman (3) has emphasized that too much tension will also hinder performance. The longer reaction times may have resulted from too much general tension, bringing the arm *above* an optimal tension for a reaction time task.

Since Freeman (2) has indicated that habituation produces a concentration of tension in the muscles involved in the task, the significant drop in reaction time during the learning tasks may be a reflec-

tion of this finding. It is possible that further studies determining the role of tension in the stream of behavior will aid in solving the problems arising from the interaction of activities.

Conclusions

The reaction time of responding to a buzzer was increased significantly when any other activity was experimentally introduced into the stream of behavior.

The buzzer-set was more affected when there was a barrier present to the immediate completion of the concomitant activity.

As completion was reached, the reaction time was significantly lowered. The set to react to the buzzer, however, did not immediately reach its optimal efficiency even when all errors were eliminated in the concomitant learning problems. As measured by the quickness of reaction there was a lag in the return to optimal efficiency of the buzzer-set.

Goal activities brought a greater interference in the buzzer-set than the non-goal activity used in this experiment. There was no significant trend in the effect upon the buzzer-set when a non-goal activity was introduced into the stream of behavior.

The greatest interference in the buzzer-set was shown in the learning problems during the first two trials when 48 percent of the errors were made.

The set to respond to an auditory stimulus was carried through varying concomitant activities without completely disappearing, though its efficiency, as measured by quickness of reaction, varied greatly with the nature of the accompanying activities. Possible interpretations in terms of the role of tension in the stream of behavior were suggested.

(Manuscript received February 12, 1945)

Bibliography

1. Daniel, R. S. The distribution of muscular action potentials during maze learning. *J. exp. Psychol.*, 1939, **24**, 621–629.
2. Freeman, G. L. The spread of neuro-muscular activity during mental work. *J. gen. Psychol.*, 1931, **5**, 479–493.
3. Freeman, G. L. The facilitative and inhibitory effects of muscular tension upon performance. *Amer. J. Psychol.*, 1933, **45**, 17–52.
4. Freeman, G. L., & Giese, W. J. The relationship between task difficulty and palmar skin resistance. *J. gen. Psychol.*, 1940, **23**, 217–220.
5. Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.
6. Pickford, R. W. An experiment on insight. *British J. Psychol.*, 1938, **28**, 412–423.
7. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt, 1938.

STANDARDIZATION OF TWO TESTS OF EQUILIBRIUM: THE RAILWALKING TEST AND THE ATAXIAGRAPH [1]

BY M. BRUCE FISHER, JAMES E. BIRREN, AND ALAN L. LEGGETT

Naval Medical Research Institute, National Naval Medical Center, Bethesda, Maryland

Measures of sensorimotor performance have often been used to determine the effects of stress conditions on the human organism (e.g., **1**, **5**, **6**, **8**, **9**). The present study represents a portion of the authors' efforts directed at the development and utilization of new and conventional psychological tests in situations wherein some indications of the effect of environmental stress are desired. The railwalking test and the ataxiagraph are two such tests that have been used to follow the performance of *S*s under conditions suspected of inducing lowered performance. This report discusses the characteristics of these tests administered to *S*s under normal circumstances.

Apparatus and Procedures

Railwalking test.—The railwalking test used in this study was a modification of an earlier railwalking test (**4**) and consisted of a narrow wooden board which the *S* had to walk, heel to toe, keeping his hands clasped behind his back. The rail was 10 feet long, one in. wide, and six in. high. White tape markers were placed along the rail at one foot intervals. Performance was scored by recording the distance in feet that the *S* successfully walked, i.e., the distance from the starting end of the board to the place where a foot was last successfully placed. The score of a trial was recorded as a whole number, fractions of a foot completed being disregarded. Trials ended when the *S* touched the floor or any part of the rail but the top to prevent falling, when he broke step or unclasped his hands to maintain balance, or when he reached the other end of the board. Clasping hands behind one's back makes balancing on the rail more difficult than being permitted to swing the arms freely. In earlier exploratory studies the *S*s were allowed to use their arms to help maintain balance, but so many perfect scores resulted as to prevent satisfactory differentiation in performance. Similarly the length of the test rail was increased from the original six feet to the present 10 feet to obtain a greater dispersion of scores.

*S*s were instructed to watch the rail several feet ahead and not to watch their feet as they walked. The ends of the rail were used alternately as the starting points for successive trials. Ten trials comprised each test session.

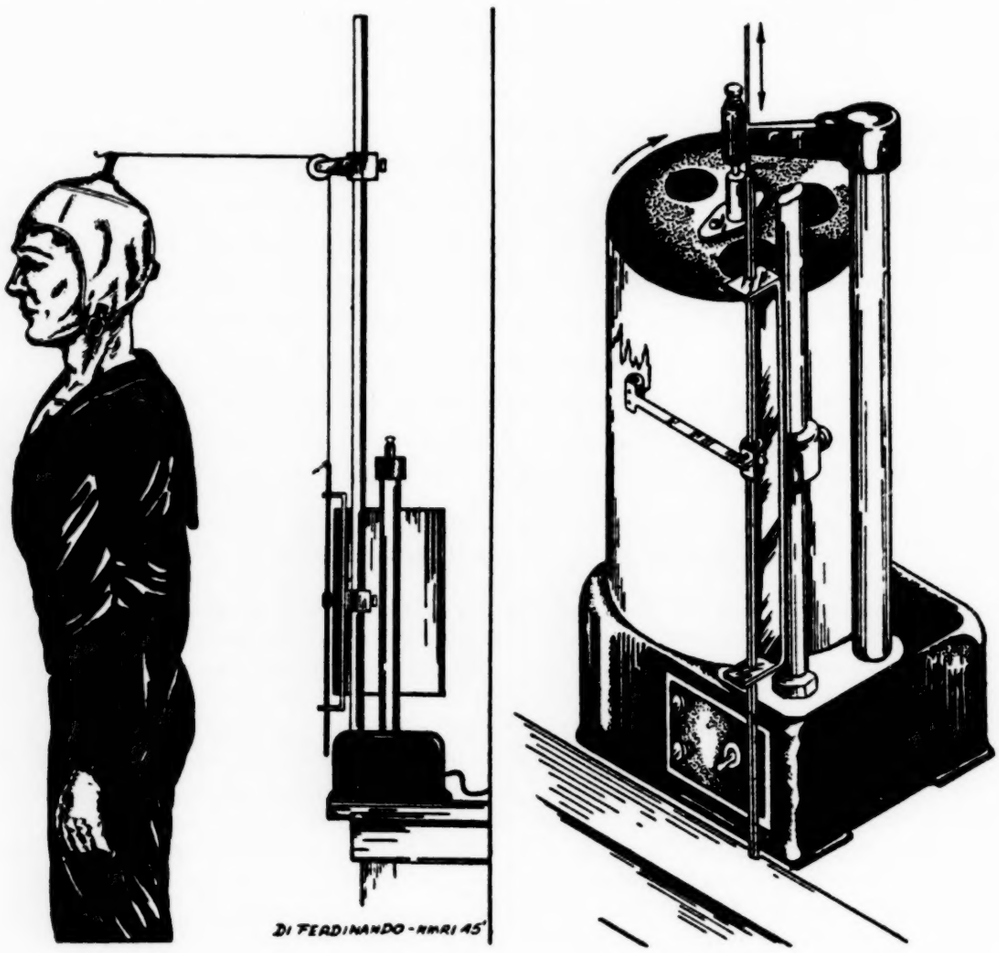
Ataxiagraph.—Body sway has been found useful as an indication of sensorimotor integrity (**1**, **2**, **3**, **7**, **8**, **9**). Measurements of body sway have ranged from the subjective evaluation of movement in the Romberg position to a variety of graphic recording techniques and the ataxiameter of Miles, a device which summated movement in the anterior-posterior and lateral planes (**7**). The ataxiagraph used in this study represents a compromise between the highly quantitative results of the ataxiameter and the qualitative advantages of graphic recording.

The technique consisted of recording the anterior-posterior movements of a person standing erect with his heels together and his feet spread in a 30° 'V' position (Fig. 1). The *S* wore a

[1] The opinions and views set forth in this article are those of the writers and are not to be considered as reflecting the policies of the Navy Department.

helmet from the top of which protruded a short rod, to which was fastened a cord which ran over a pulley and translated the *S*'s movement into equivalent vertical movements of an inkwriter on a kymograph. The weight of the inkwriting assembly was 59.9 gm., a perceptible but constant amount.

Two test lengths have been used: (*a*) an eight-min. test of four two-min. trials, with eyes alternately open and closed, beginning with eyes open, (*b*) a four-min. test of two two-min. trials, the first with eyes open, the second with eyes closed. *S*s were instructed to stand erect without straining during a trial. A 30-sec. rest period was allowed between trials, during which the cord was released from the helmet and the *S*s were allowed to move freely. There was a clear space



A *B*

FIG. 1. Ataxiagraph, a kymograph modified to record body sway in one plane. *A*. Subject in position for testing. *B*. Detail of kymograph.

without furniture or other objects at least 10 feet by 5 feet immediately in front of the *S*'s position to reduce visual cues of body movement during the eyes-open trials.

Scoring of the graphic record was achieved by measuring the length of the line traced by the *S*'s movements with a Dietzgen map measure and subtracting the horizontal distance traveled by the kymograph surface during the trial. The result was then converted into cm. of anterior-posterior movement by the equation:

$$\text{Score} = 1.128\,[(\text{length of line}) - (\text{horizontal distance})] + 4.003$$

This formula is the least-squares fit of 98 measured excursion scores and map measure scores. The excursion scores for this calculation were secured by careful measurement of the vertical distances between successive maxima and minima on 56 one-min. records and 42 two-min. records.

These measurements were accurate to half a mm. The 56 one-min. records were unselected and in this sample, in which the excursion score range was 6.4–22.5 cm., the two scores had a product-moment correlation of 0.94 (σ_{est} = 1.34 cm.). The 42 two-min. records were selected to cover the range of scores to be expected in both normal and pathological populations. The standard error of estimate for the total of the 98 scores on the basis of the line of best fit given above is 3.96 cm. This group included a range of scores from 6.4 to 305.4 cm. In the range of normal Ss the standard error was 2.9 cm. for a single trial.

The use of the above formula manifestly violates the Pythagorean theorem by failing to square the dimensions entering into it. That this error is small is indicated by the standard errors of estimate. Scoring with the map measure, which takes less than a fourth the time required for the exact measurement, is therefore justified and has been used routinely.

Results

Railwalking.—The distribution of scores on the railwalking test is given in Table I. It has a slight positive skew [2] (Sk = .05) but does not differ significantly, for a sample this size, from a 'normal' curve.

TABLE I

Distribution of Railwalking Scores

Score Intervals (feet)	Frequencies
97–100	–
94– 96	–
91– 93	1
88– 90	0
85– 87	0
82– 84	0
79– 81	1
76– 78	0
73– 75	1
70– 72	1
67– 69	1
64– 66	0
61– 63	2
58– 60	5
55– 57	4
52– 54	13
49– 51	3
46– 48	10
43– 45	8
40– 42	10
37– 39	14
34– 36	7
31– 33	8
28– 30	16
25– 27	16
22– 24	10
19– 21	9
16– 18	6
13– 15	5
10– 12	–

N = 151
Mean = 37.12
Median = 36.36
σ = 14.45

[2] $Sk = \frac{3(\bar{X} - \text{med.})}{\sigma}$.

The split-half reliability (odd versus even trials) of this test was 0.77 (N = 151) for the first test session, and 0.85 (N = 78) for the second administration, both reliabilities being corrected for test length by the Spearman-Brown prophecy formula. A retest reliability of 0.67 was found for 78 *Ss* retested within one week.

When 18 *Ss* were run on the railwalking test, twice a day, for four days, the group means evidenced regular improvement. The practice that the *Ss* received was not spaced to show in detail the learning

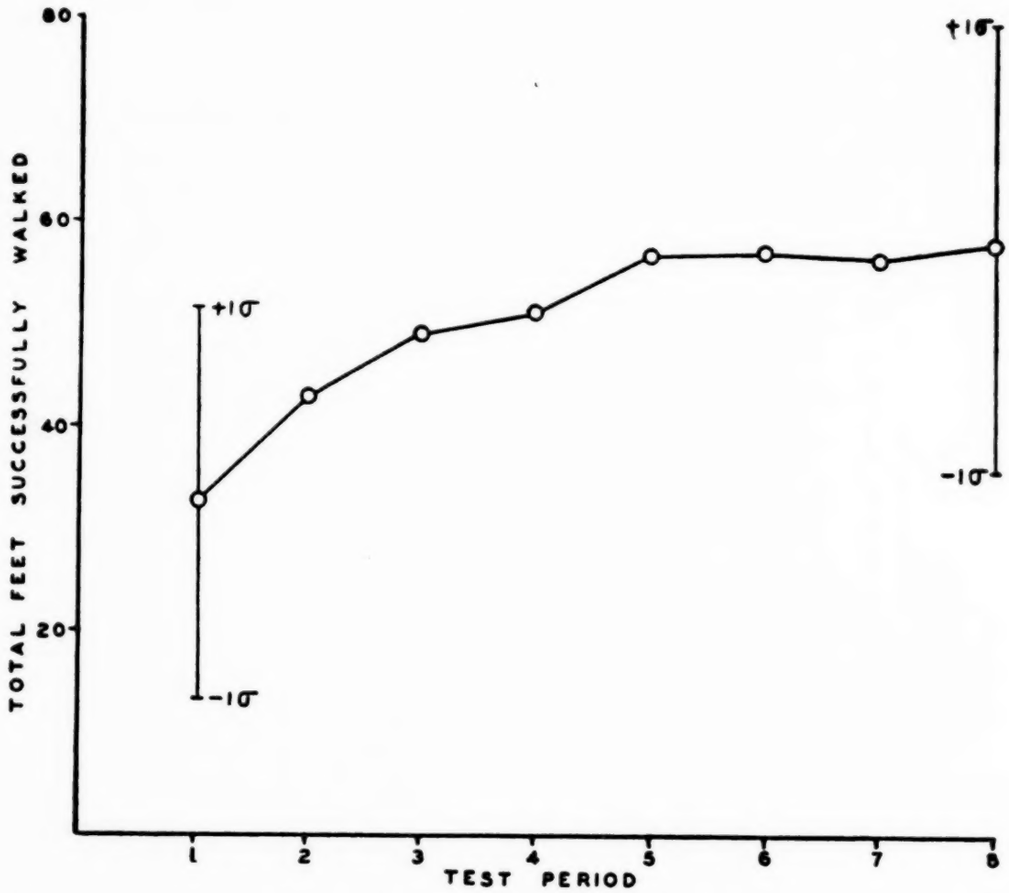


FIG. 2. Learning on the railwalking test. The graph shows the mean scores of 18 *Ss* who were tested twice a day for four days. The vertical bars represent the size of the standard deviation of the scores in the first and eighth test periods.

phenomena of the test, but rather to meet the practical requirements of a specific testing situation. The *Ss* were given their two daily tests with only a 10-min. rest between, the pairs of daily tests being 24 hours apart. Even though the two daily test sessions were so close together, the graph of the data (Fig. 2) does not reflect any inhibition or reminiscence phenomena to qualify the conclusion that the test scores regularly improve with practice. The improvement in performance between the mean scores of the first and the eighth tests is 78 percent.

Ataxiagraph.—Frequency distributions of ataxiagraph scores for both the four-min. test and the eight-min. test are given in Table II. It is apparent that there is considerable skewing in the data. For the two distributions *Sk* is 0.86 and 0.63, respectively. Skewness to this

TABLE II

DISTRIBUTIONS OF ATAXIAGRAPH SCORES

Score Intervals (cm./4 min.)	Frequencies: 4-min. Test (2 min. eyes open plus 2 min. eyes closed)	8-min. Test (mean of two 4-min. tests)
0– 5	–	–
5– 10	–	–
10– 15	–	–
15– 20	1	1
20– 25	2	2
25– 30	1	2
30– 35	8	2
35– 40	5	8
40– 45	4	11
45– 50	18	13
50– 55	15	15
55– 60	9	9
60– 65	10	5
65– 70	7	12
70– 75	9	4
75– 80	5	11
80– 85	7	4
85– 90	6	3
90– 95	5	4
95–100	3	7
100–105	2	2
105–110	4	3
110–115	2	3
115–120	1	3
120–125	3	0
125–130	0	2
130–135	0	2
135–140	1	1
140–145	2	1
145–150	1	0
150–155	0	0
155–160	0	0
160–165	0	0
165–170	1	2
170–175	0	0
175–180	–	1
N =	132	133
Mean =	67.75	70.02
Median =	61.5	63.5
σ =	21.84	30.82

extent cannot be dismissed as a sampling error since both measures are more than three times the standard error of χ.

Reliability coefficients were determined for ataxiagraph scores for both the eyes-open and eyes-closed conditions by correlating the pairs of two-min. trials for each from the eight-min. tests. The reliability of the mean eyes-open score is 0.89 and of the mean eyes-closed score,

also 0.89 (N = 120). Both of these correlation coefficients were corrected for test length by the Spearman-Brown prophecy formula.

A close relation existed between the eyes-open and eyes-closed scores in the four-trial (eight-min.) test. The correlation coefficient was 0.77 between the first eyes-open and first eyes-closed scores, and 0.81 between the second eyes-open and second eyes-closed scores (N = 132). This relation was accepted as sufficient justification for combining the eyes-open and eyes-closed scores into a single body-sway score. Evidence that the procedure was sound was furnished by the high correlation between the sums of the first eyes-open and eyes-closed scores, and the second eyes-open and eyes-closed scores, the coefficient being 0.92. If this is corrected for test length, the resulting reliability coefficient for the eight-min. test is 0.96 (N = 132). The reliability of the four-min. test is arrived at by correcting 0.77 for test length, which becomes 0.87.

This correlation is significant in view of the increase in body sway upon closing the eyes. The mean score, eyes open, is 28.5 cm. per two min.; with eyes closed, it is 41.6 cm. per two min. (N = 133). The difference between these means is highly significant statistically.

The retest stability of the ataxiagraph scores was determined by retesting 85 of the *Ss* within one to six days after the initial test, the retest reliability coefficient being 0.87. It was not possible to maintain a constant interval between the two tests, although most of the retest sessions occurred within two days of the original measurements. No attempt was made to control time of day in testing, since unpublished analysis of other sensorimotor test data has shown little diurnal fluctuation.

Practice in taking the test does not result in systematically improved ataxiagraph performance. No regular decline in scores was noted when the means of 19 *Ss* were followed for 10 measurements covering a period of one month. This finding is in general agreement with that of other investigators (1, 4).

When the ataxiagraph scores of 43 *Ss* were correlated with their railwalking scores the correlation was found not to differ significantly from zero (r = .02 ± .15).

Discussion

The two tests described in this report are apparently quite different in the kinds of ability required for good performance. On *a priori* judgment it would appear that both tests involve aspects of equilibrium, one static and the other dynamic. Both the zero correlation between ataxiagraph scores and railwalking scores, and the difference in the effect of practice on test scores, are evidence that

quite different sensorimotor functions are being measured. In contrast to the marked improvement in railwalking with practice, little improvement in body sway was noted. This small improvement in body sway appears as an indication that maintenance of stable erect posture is probably a phylogenetically more primitive mechanism. In view of the evidence, however, that attention and other cortical phenomena modify to some extent the amount of body sway, it would be fallacious to assume that static equilibrium is wholly uninfluenced by the motor cortex. It is anticipated that future investigations will not only study the effects of environmental stress on such sensorimotor tests, but may also uncover the basis of these equilibratory functions. Analysis of the relations of Barany chair reactions (fall and nystagmus) to body sway and railwalking performance will be one approach to the significant components in the maintenance of postural equilibrium.

Summary

1. One hundred and fifty-one male naval enlisted personnel were given 10 trials on a railwalking test. The object of this test is to walk the length of a rail which is one in. wide and 10 feet long, placing heel to toe, with hands clasped behind the back. Performance was scored as the sum of the distances walked in 10 trials, the maximum score being 100. The odd-even reliability coefficient of the first test on the 151 *Ss* was 0.77, corrected for test length. The odd-even reliability of a second test given to 78 of these *Ss* was 0.85, corrected for test length.

2. The correlation between the first and second tests (within one week) for the 78 *Ss* was 0.67.

3. One hundred and thirty-three male naval enlisted personnel were used as *Ss* to record body sway on an ataxiagraph, an inkwriting kymograph arrangement for recording anterior-posterior body sway. Four two-min. trials were given, two with the *S*'s eyes open and two with his eyes closed, his feet being in the 30° 'V' position. The conditions were run alternately, with a half-min. rest between trials. Scores were recorded as cm. of anterior-posterior movement in two min., as determined from the length of the line recording this movement.

4. The correlation coefficient between the first eyes-open and first eyes-closed score was 0.77; between the second eyes-open and second eyes-closed score the coefficient was 0.89. The sum of the first eyes-open and eyes-closed scores, correlated with the sum of the second eyes-open and eyes-closed scores, yielded a coefficient of 0.92. The split-half reliability of the four-min. test, corrected for length, is,

thus, 0.87; the split-half reliability computed for the eight-min. test is 0.96 ($N = 133$).

5. Eighty-five of the 133 *Ss* were retested within one week of the first test session to study the stability of the scores. The retest reliability of the eight-min. test was 0.87.

6. Forty-three *Ss* were tested on both the ataxiagraph and railwalking tests. The correlation of railwalking score with ataxiagraph score was not significantly greater than zero.

7. Marked practice effects were noted on the railwalking test in 18 *Ss* who were tested twice a day for four days. The improvement between mean scores on the first and eighth tests was 78 percent.

8. Little improvement in ataxiagraph scores was observed in 19 *Ss* who were tested 10 times over a one-month period.

Conclusions

Both test procedures reported in this paper appear suitable for inclusion in a group of tests to measure sensorimotor performance. Of the two tests, the ataxiagraph yields the more reliable scores. The reliability of the railwalking test cannot be regarded as entirely satisfactory for work with small populations. Improvement could be expected from doubling the test length. The estimated reliability of a 20-trial test, based on the reliability of the first 10 trials, is 0.87.

No relation exists between railwalking and ataxiagraph scores; performance on the two tests, therefore, is presumed to depend on different factors. A close relation exists between body sway in the eyes-open and eyes-closed conditions. Since the same function is apparently measured with the eyes either open or closed, it is considered feasible to use only one trial with eyes open and one trial with eyes closed in routine testing. The sum (or the mean) of the scores of these two conditions can be regarded as a general measure of body sway. The use of both eyes-open and eyes-closed trials rather than four minutes of either condition alone has the advantage of making possible the detection of unusual influences which might affect the two scores differentially. This length of test gives satisfactory reliability for most purposes, although the eight-min. test may be resorted to where greater precision is indicated.

(Manuscript received February 15, 1945)

References

1. Edwards, A. S. Effects of the loss of one hundred hours of sleep. *Amer. J. Psychol.*, 1941, 54, 80–91.
2. Fearing, F. S. The factors influencing static equilibrium. I. *J. comp. Psychol.*, 1924, 4, 91–121.

3. FEARING, F. S. The factors influencing static equilibrium. II. *J. comp. Psychol.*, 1924, 4, 163–183.
4. HEATH, S. R. The military use of the railwalking test as an index of locomotor coordination. *Psychol. Bull.*, 1943, 40, 282–284.
5. MALMO, R. B., & FINAN, J. L. A comparative study of eight tests in the decompression chamber. *Amer. J. Psychol.*, 1944, 57, 389–405.
6. MCFARLAND, R. A. The psychological effects of oxygen deprivation (anoxemia) on human behavior. *Arch. Psychol.*, 1932, No. 145.
7. MILES, W. R. Static equilibrium as a useful test of motor control. *J. indust. Hyg.*, 1922, 3, 316–331.
8. MILES, W. R. *Alcohol and human efficiency.* Washington, D. C. Carnegie Institution, 1924.
9. *Fatigue and hours of service of interstate truck drivers.* Public Health Bulletin 265. Federal Security Administration, U. S. Public Health Service, Washington, D. C., U. S. Gov. Printing Office, 1941.

THE CONSTRUCTION OF A PUNCH MAZE

BY WESTON A. BOUSFIELD

University of Connecticut

AND

VINCENT NOWLIS

Indiana University

The writers have developed the maze herein described for use in classes where the schedule calls for a simply administered and easily scored experiment on learning. The principle of this type of maze allows its construction in a variety of forms, some of which may well prove useful in research. Considerations governing the design of this model were economy of time and materials, and construction of the units by relatively unskilled student help.

The maze is made of two pieces of heavy cardboard perforated as indicated in Figs. 1-*A* and 1-*B*, and hinged together by means of gummed cloth mending tape as shown in Fig. 1-*C*. The record sheet is inserted between the hinged covers. The circular perforations of the front cover, Fig. 1-*A*, are 1/8 in. in diameter and are made with a small paper punch. This cover contains, in addition to the single upper and lower holes (the start and finish of the maze), 15 pairs of holes arranged in parallel columns. The 17 holes of the back cover are 3/16 in. in diameter, and define the start and finish of the maze and the 15 correct right-left choices. In using the maze, the record sheet is inserted between the hinged covers, and the *S* is provided with a three-in. nail to serve as a fairly blunt stylus. The *S*'s task is to learn the sequence of 15 right-left choices. Starting from the upper single hole, the *S* chooses between the right and the left of the first pair of holes. If the choice is correct, the stylus easily penetrates the inserted paper because of the hole in the back cover and a small hole is left in the record sheet. On the other hand, if the choice is incorrect, the stylus will not penetrate the paper, and leaves only an indentation which may be easily distinguished from a hole. The *S* immediately corrects his error and then proceeds to the next pair of holes where he again makes a right-left choice. In this way the *S* makes a permanent record of his trials showing the location of all errors. Scoring is made easier by instructing the *S* to punch vigorously, and by using a record sheet of the type illustrated in Fig. 1-*C*. After each 15-choice trial the *S* moves this sheet to the left so as to bring the next section be-

neath the holes of the maze cover. It may be noted that the record sheet allows for 10 trials and has spaces at the bottom for indicating the total errors for each trial and spaces at the side for the total errors for each of the 15 choice-points. A set of holes is provided for preliminary practice, in which the *S* gets the feel of the stylus as it makes holes (correct choices) and indentations (incorrect).

Several details concerning the construction of the apparatus merit attention. The covers should be hinged as carefully as possible so as to bring the holes into alignment. Wire fasteners are punched

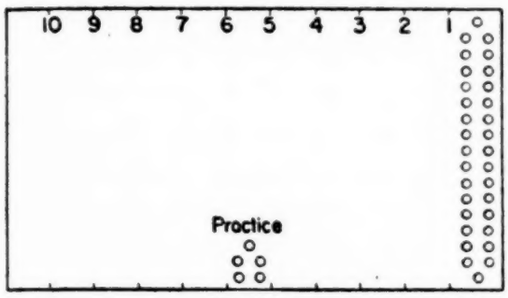


FIG. 1-A

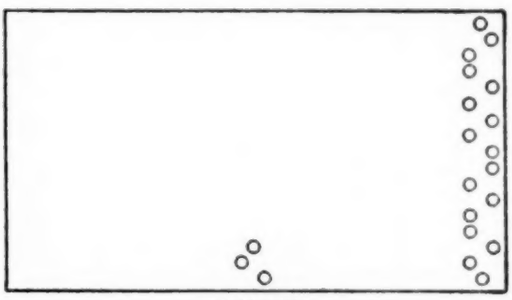


FIG. 1-B

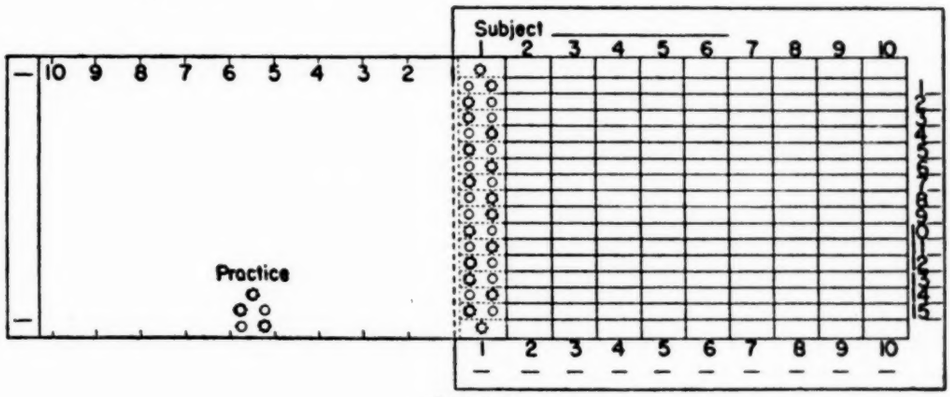


FIG. 1-C

FIGURES 1-*A*, 1-*B*, AND 1-*C*

FIG. 1-*A*. Front cover of the punch maze.

FIG. 1-*B*. Back cover of the punch maze.

FIG. 1-*C*. Assembled punch maze with record sheet in position for first trial.

(Dimensions: 1-*A* and 1-*B*, 5¾ × 11 in.; holes in 1-*A*, *not drawn to scale*, ⅛ in. in diameter; holes in 1-*B*, *not drawn to scale*, 3/16 in. in diameter; record sheet 8½ × 11 in.)

through the binding tape and individual covers to assure holding of the tape. There is some advantage to be gained by punching the holes of the back cover after the two covers have been hinged. The procedure in this case is to draw circles on the back cover by inserting the pencil point through the holes of the front cover. These circles serve as a guide for punching the lower holes which are made larger than the upper ones to compensate for any failure of alignment of the covers as the apparatus is used.

In using the maze we have employed the technique of having the

Ss make their choices in time with a buzzer sounding at intervals of two sec. This assures a constant total time for each trial. A summary of data for 100 *Ss* appears in Tables I and II. These data were obtained in laboratory sections of 18 to 25 students in an introductory course in psychology. The criterion of mastery of the maze was three

TABLE I

ERRORS PER TRIAL FOR FIRST 10 TRIALS OF PUNCH MAZE MADE BY 100 *Ss*

Errors	Trials									
	1	2	3	4	5	6	7	8	9	10
0		2	9	20	28	31	48	52	62	64
1		5	13	12	15	21	21	17	17	14
2	1	9	14	20	15	15	15	9	12	12
3	4	14	20	17	14	15	7	14	5	6
4	10	16	13	11	14	11	5	5	3	4
5	18	13	7	7	7	4	2	0	0	
6	10	11	9	9	5	1	1	2	1	
7	9	16	8	2	0	2	1	0		
8	14	4	3	2	1			0		
9	9	4	2		1			1		
10	9	1	1							
11	5	1	0							
12	1	1	0							
13	2	0	1							
14	3	2	0							
15	5	1	0							
Mean Errors	7.60	5.16	3.63	2.66	2.25	1.80	1.15	1.18	.74	.72

TABLE II

ERRORS AT SUCCESSIVE CHOICE POINTS FOR FIRST 10 TRIALS OF PUNCH MAZE MADE BY 100 *Ss*

Errors	Choice Points														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	30	29	12	40	35	30	14	12	10	48	26	20	8	44	44
1	24	43	18	33	30	32	18	12	17	31	28	33	22	29	39
2	19	13	13	24	20	20	17	22	22	13	17	21	24	21	11
3	11	6	21	2	11	10	19	18	18	6	15	16	17	5	6
4	8	5	15	1	3	3	19	13	13	2	9	5	7	0	
5	4	1	9		1	0	4	8	8		3	2	14	1	
6	2	2	5			3	3	9	5		0	2	2		
7	2	1	2			2	4	4	4		1	0	3		
8			2				1	2	1		1	0	3		
9			3				1		1			1			
10									1						
11															
12															
13															
14															
15															
Mean Errors	1.73	1.31	2.99	.91	1.20	1.46	2.68	3.00	2.92	.83	1.73	1.74	2.76	.91	.79

errorless trials. When this criterion was attained by the ninth or an earlier trial, the *Ss* received credit for absence of errors on subsequent trials even though they did not continue with the maze. The instructor acted as *E* and each student worked with his own maze. With respect to classroom procedure we have found it expedient to supplement verbal instructions with a demonstration of the manipulation of the maze. A further detail of technique is to instruct the *Ss* to shift the record sheet to the next section immediately following completion of a trial. This prevents study of the pattern of holes in the record sheet during the intervals between successive trials.

Through repeated use of the maze the holes in the front cover defining the correct choices become marked to the extent that they provide cues. When this happens, these cues may be eliminated by putting a strip of gummed cloth binding tape over the holes. The tape is then punched from the reverse side of the cover by inserting the punch through the original holes.

(Manuscript received February 3, 1945)